The Biometrics Section of the American Statistical Association

BIOMETRICS

TABLE OF CONTENTS

The Analysis of Covariance . . . . D. B. DELURY 153

An Experimental Design Used to Estimate the Optimum Planting Date for Cotton . . R. L. ANDERSON AND H. L. MANNING 171

On the Use of Parallel or Nonparallel Systems of Transformed Curves in Bioassay: Illustration in the Quantitative Complement-Fixation Test . . . . . . WILLIAM R. THOMPSON 197

Queries 211

The Biometric Society 216

News and Notes 220

Number 3 September 1948 Volume 4

Material for *Biometrics* should be addressed to the Chairman of the Editorial Committee, Institute of Statistics, North Carolina State College, Raleigh, N. C.; and material for Queries should go to "Queries", Statistical Laboratory, Iowa State College, Ames, Iowa, or to any member of the committee.

Officers
American Statistical Association

President, George W. Snedecor; Directors, Isador Lubin, Lowell J. Reed, Walter A. Shewhart, Samuel A. Stouffer, Frederick F. Stephan, Willard L. Thorp; Vice-Presidents, A. Ford Hinrichs, Rensis Likert, Abraham Wald, Harry R. Wellman; Secretary-Treasurer and Managing Editor of Publications, Merrill M. Flood.

Officers
Biometrics Section

Chairman, Joseph Berkson; Secretary, John H. Watkins; Section Committee, D. J. Finney, Margaret Merrell, P. J. Rulon and J. W. Tukey.

Editoral Committee
Biometrics

Chairman, Gertrude M. Cox; Members, C. I. Bliss, W. G. Cochran, Churchill Eisenhart, J. W. Fertig, H. C. Fryer, Horace Norton, A. M. Mood, G. W. Snedecor, and Jane Worcester.

Membership dues in the American Statistical Association are $8.00 a year, of which $4.00 is for a year's subscription to the quarterly *Journal*, $.75 is for a year's subscription to *The American Statistician*. Members of the Association may subscribe to *Biometrics* at a special rate of $2.00. Dues for associate members of the Biometrics Section are $3.00 a year, of which $2.00 is for a year's subscription to *Biometrics*. Single copies of *Biometrics* are $1.00 each and annual subscriptions are $3.00. Subscriptions and applications for membership should be sent to the American Statistical Association, 1603 K Street, N.W., Washington 6, D. C.

Entered as second-class matter, May 25, 1945, at the post office at Washington, D. C., under the Act of March 3, 1879. Biometrics is published four times a year—in March, June, September and December—by the American Statistical Association for its Biometrics Section. Editorial Office: 1603 K Street, N.W., Washington 6, D.C.

THE ANALYSIS OF COVARIANCE*

D. B. DeLury

Virginia Polytechnic Institute
and the Ontario Research Foundation

SECTION I

THE WHOLE OF this discussion is based on the data of a single experiment, the details of which have been published under the title "The Effect of Atropine and Quinidine Sulphate on Atrophy and Fibrillation in Denervated Skeletal Muscle." [1]

For the present, we may adopt the view that the experiment was conducted to compare the effects of four different drugs in delaying the atrophy of denervated muscles. The structure of the experiment was, briefly, as follows. A number of rats were put randomly into four groups; a certain muscle in the hind leg of each rat was deprived of its nerve supply by severing the appropriate nerves, the leg to be denervated (right or left) being chosen at random. Each of the four groups was assigned to treatment by one of the drugs and the treatments were continued throughout the course of the experiment. Four days after treatment was begun, four rats were chosen randomly from each of the four groups and measures of atrophy were obtained from them. This procedure was repeated after 8 and 12 days.

Atrophy is usually measured by the loss of weight of the muscle after denervation. The weight of the muscle can be determined only by killing the animal and removing and weighing the muscle; hence the weight of the muscle at the beginning of the experiment is not known and a direct measure of weight loss is not obtainable. The device ordinarily used to circumvent this difficulty is to obtain the weight of the same muscle from the other leg (which was not denervated), at the same time as the denervated muscle is taken, and to assume that the weight of the intact muscle, at the end of the experiment, is the same as that of the denervated muscle at the beginning of the experiment.

---

*A revision of an expository paper given at a joint meeting of the Biometrics Section and the Institute of Mathematical Statistics, held in conjunction with the 113th Annual Meeting of the American Association for the Advancement of Science, Boston, Massachusetts, December 28, 1946. Section I contains the substance of the paper as presented. Section II is based on comments given by Dr. C. I. Bliss in a prepared discussion. Section III has been added later.

TABLE I

Days	Drug	body weight		muscle weight		fibrillation
		initial x	final y	denervated u	intact v	w
4	A large atropine	217	196	0.94	1.16	17.5
		246	218	1.16	1.20	10
		256	216	1.26	1.31	12
		200	165	0.85	1.01	14
4	B moderate quinidine	198	202	1.19	1.25	8.5
		248	231	1.15	1.21	7
		180	187	0.86	1.16	17.5
		218	230	1.21	1.38	16.5
4	C moderate atrophine	264	231	1.22	1.34	10
		200	170	0.90	1.00	12
		210	189	1.00	1.03	7
		192	185	1.00	1.14	14.5
4	D saline	181	193	0.99	1.17	12
		266	285	1.51	1.73	14
		274	266	1.55	1.75	17.5
		180	188	0.98	1.15	12.5
8	A	265	183	0.91	0.91	5
		248	190	0.73	0.89	7
		238	166	0.52	0.77	14
		180	169	0.65	0.97	13
8	B	186	200	0.87	1.24	6.5
		220	221	1.04	1.42	9
		199	230	0.88	1.40	11
		240	246	0.96	1.38	9
8	C	178	162	0.67	0.87	4
		188	181	0.72	1.00	9
		250	235	1.08	1.34	9
		195	182	0.75	1.04	10
8	D	194	207	0.97	1.39	6.5
		274	267	1.07	1.76	14
		222	237	1.16	1.66	15
		274	243	1.04	1.69	13
12	A	198	165	0.34	0.70	5
		175	150	0.43	0.60	5
		199	159	0.41	0.66	15
		224	163	0.48	0.60	14
12	B	233	242	0.41	1.03	11
		250	226	0.87	1.30	3
		289	300	0.91	1.67	15
		255	252	0.87	1.52	11
12	C	204	181	0.57	0.97	16
		234	181	0.80	1.10	13
		211	180	0.69	0.87	22
		214	200	0.84	1.22	15
12	D	186	243	0.81	1.61	4
		286	297	1.01	1.80	4
		245	264	0.97	1.68	7
		215	228	0.87	1.60	6

We might expect that the difference between these two numbers (intact minus denervated) would be used in whatever analysis is undertaken to compare the effects of the four drugs. However, it seems to be generally accepted that the *ratio*, (denervated)/(intact), is the proper combination of these two numbers to use.

Table I contains the measurements yielded by the experiment. For the moment, we are concerned only with the weights of the denervated and intact muscles. The ratios (total denervated)/(total intact), and the differences, for the 12 sets of 4 rats, are listed in Table II.

TABLE II

	Ratios				Differences			
	A	B	C	D	A	B	C	D
4 days	.90	.88	.91	.87	.47	.57	.39	.77
8 days	.79	.69	.76	.65	.63	.69	1.03	1.24
12 days	.65	.55	.70	.55	.90	2.46	1.26	3.03

Putting aside the question of a test of significance, it would appear that responses to the drugs are of two kinds; A and C are alike, B and D are alike, but there is a perceptible difference between the two pairs. This decision is reached on the basis of either the proportions or the differences.

Now drug D was, in fact, simply a saline solution and could have had no effect on atrophy; drug B was a moderate dose of quinidine sulphate, which might have acted to delay the deterioration of the denervated muscles, but apparently did not to any considerable degree; drugs C and A were moderate and large doses of atropine sulphate which, we might conclude, did exert some beneficial influence in retarding the progress of atrophy.

Altogether this makes a not unreasonable picture. However, let us inquire more closely into the numbers which have gone into this simple analysis. Why should the ratio (denervated)/(intact) be used? Is there any reason why we should not conduct the analysis on the weights of the denervated muscles alone?

The answer to the second of these questions is obvious: an analysis based on the weights of the denervated muscles would be legitimate, but would probably be very insensitive, because differences among the initial weights of these muscles would enter into the analysis as experimental error and might well obscure real effects. Furthermore, we must conclude that the weight of the intact muscle is introduced in the ratio (denervated)/(intact) in the hope of preventing these differences among initial muscle weights from inflating the experimental error. It should be clear, however, that this device is very unlikely to succeed, *even if the intact muscle weight were identical with the initial weight of the denervated muscle*, because to suppose that the ratio (denervated)/(initial) is independent of the initial weight is equivalent to assuming that the amount of atrophy is proportional to the initial weight. The constant of proportionality could, of course, change from one treatment to another, but in any case this assumption is one that rarely is met in practice. Likewise, if we were to base an analysis on the difference, (initial) − (denervated), in the expectation of removing the effects of variation among the initial weights, we would be successful only if the amount of atrophy were independent of the initial weight. This assumption is not appreciably better than that of proportionality.

When, to these weaknesses, is added the obvious fact that the final weight of the intact muscle is certain to differ from the initial weight of the denervated muscle, the use of the intact muscle weight, in either a proportion or a difference, is seen to be open to serious objection.

Of the many possible reasons why the final intact muscle weight is likely to differ from the initial weight of the denervated muscle, two deserve special mention.

(1) Throughout the course of the experiment, the animal may grow normally, or nearly so, apart from the denervated muscle. The intact muscle is therefore heavier at the end of the experiment than it was at the beginning. This does not, in itself, render the intact muscle undesirable for use in the analysis. Indeed, the effect may be entirely good, because the intact muscle weight would reflect the weight that the denervated muscle would have attained if the nerve had not been cut.

(2) The drugs employed may affect directly the weight of the intact muscle. If this is the case, the consequences of using this number to indicate the initial weight of the denervated muscle are devastating. It was in anticipation of this possibility that records of the initial and final weights of the animals were kept in this experiment.

Only a glance at Table I is needed to see that the drugs did affect the weights of the intact muscles. The totals of the weights of both intact and denervated muscles are shown in Table III. (From this point on, the decimals are dropped from the observations on u and v.)

Recalling that, on the basis of either ratios or differences, it was decided that treatments A and C might be producing beneficial effects and that B and D were not, it is somewhat startling to observe that the weights of the denervated muscles are *lower* in the A and C groups than in the B and D groups. It is seen too that the same thing is true of the intact muscle weights. Now we can see clearly what has happened. The ratios computed for the A and C drugs were high, not because their numerators were *large* (the reverse was the case) but because their denominators were *small*. Similar remarks apply to the differences. We see also that the conclusions reached by means of ratios and differences are exactly the reverse of those which ought to be reached (assuming for the moment that *some* conclusion is warranted).

TABLE III

	A		B		C		D	
	den. u	int. v	den. u	int. v	den. u	int. v	den. u	int. v
4 days	421	468	441	500	412	451	503	580
8 days	281	354	375	544	322	425	424	650
12 days	166	256	306	552	290	416	366	669

It should be clear at this point that if we are to make allowance in our analysis for the variation among the initial weights of the denervated muscles, we must have either their values or measurements on some variable which follows their values as closely as possible and which is not subject to the influence of the drugs. The initial weights we cannot obtain, but it seems reasonable to suppose that these initial weights are closely related to the total weights of the animals at the beginning of the experiment. This supposition should, of course, be checked against experimental evidence. The technique employed to perform this check and to make proper allowance in the analysis for variation among the initial weights is known as The Analysis of Covariance.

The analysis of covariance is essentially an application of standard regression theory to the problem at hand. It seems obvious that the dependence of the final weight of the denervated muscle on its initial weight, or on some variable which is correlated with it, is likely to lie somewhere between the extremes of proportionality and additivity which were examined earlier. Clearly it is necessary that we assess this dependence from the observations of our experiment, and this is precisely the job for which the theory of regression was developed. This situation differs from the general regression problem only through being somewhat simpler. The simplicity derives from the fact that the observations to which regression methods are to be applied come from a balanced experiment.

The computations required are exhibited in Tables IV and V. An ordinary analysis of variance is performed on the variables x (initial total body weight) and u (final denervated muscle weight). The results of these computations are entered in the first four rows of Table IV, in the columns labelled (x^2) and (u^2). The column headed (xu) gives the breakdown of the sum of the products of x and u, according to the same rules as are used in separating the sums of squares of x and of u.

TABLE IV
SUMS OF SQUARES AND PRODUCTS

row		d.f.	(x^2)	(xu)	(u^2)
1	times	2	264	−1736	13269
2	drugs	3	2969	3451	7931
3	times × drugs	6	9053	2083	1529
4	error	36	38244	14562	10111
5	times + error	38	38508	12826	23380
6	drugs + error	39	41213	18013	18042
7	times × drugs + error	42	47297	16645	11640

TABLE V
SUMS OF SQUARES FOR REGRESSION AND DEVIATIONS FROM REGRESSION

row		regression		deviations		
		d.f.	s.s.	d.f.	s.s.	m.s.
1	error	1	5545	35	4566	130
2	times + error	1	4272	37	19108	
3	drugs + error	1	7873	38	10169	
4	times × drugs + error	1	5858	41	5782	
5	times			2	14542	7271
6	drugs			3	5603	1868
7	times × drugs			6	1216	203

Only the simplest arithmetic is involved in passing from Table IV to Table V. The s.s. in the error row of Table V are obtained by combining the numbers in the error row in Table IV; the s.s. for regression is

$(14562)^2/(38244)$ and the s.s. for deviations from regression is $10111 - (14562)^2/(38244)$. The rows for (times + error), (drugs + error) and (times × drugs + error) in Table V are computed according to the same rule from the corresponding rows of Table IV. The entries in rows 5, 6 and 7 of Table V are obtained by subtracting the numbers in row 1 from those in rows 2, 3, 4.

Underlying this calculation there is envisaged a regression line, $U - \bar{u} = b(x - \bar{x})$, which expresses the dependence of u on x. This line is fitted using the sums of squares and products from the error row of Table IV, in the expectation that any dependence detected here will not be reflecting time or drug effects. The slope b of the line is therefore $(14562)/(38244) = .3808$.

This line can be used to compare two observations or two averages of u, undisturbed by differences between their corresponding x-values, provided the line expresses adequately the dependence of u on x. If (u_1x_1), (u_2x_2) are two pairs of observations or averages, it is seen that u_1 and u_2 may be expected to differ by $b(x_1 - x_2)$, by reason of the difference between their x-values and allowance should be made for this fact in trying to interpret the difference $(u_1 - u_2)$. In practice, instead of evaluating each pair separately, it is simpler to reduce each (u, x) pair to an equivalent pair $(u', \bar{x})$, where $\bar{x}$ is the grand mean of x. Differences among the u'-values will then be unaffected by differences among the x's.

Totals of x and u, taken from Table I, and the adjusted values u' are given in Table VI.

TABLE VI

DENERVATED MUSCLE WEIGHTS, ADJUSTED FOR INITIAL BODY WEIGHT

	A			B			C			D		
	x	u	u'	x	u	u'	x	u	u'	x	u	u'
4 days	919	421	411	844	441	459	866	412	422	901	503	499
8 days	931	281	266	845	375	393	811	322	353	964	424	396
12 days	796	166	202	1027	306	254	863	290	301	932	366	351

The manner in which the adjusted values u' vary from treatment to treatment points to substantial differences among the responses to the four drugs. We may well feel that a test of significance for these differences is required here. The numbers in Table V were calculated to provide such tests. The mean squares in the last column of Table V may properly be compared in an F-ratio. Thus F = (1868)/(130) =

14.4, with 3 and 35 d.f., tests the drug mean square against that for error. The mean squares for time and time × drug may be tested in the same way.

A test of significance for b seems to be not particularly necessary. After all, the variable x was selected with considerable a priori assurance that u would be closely related to it. However, such a test, if desired, is readily obtained from numbers already calculated. In the error row of Table V, the 1 d.f. for regression exhibits the amount by which the "total" error s.s. is diminished by introducing the regression of u on x; a comparison of this with the residual s.s. indicates the effectiveness of the regression in reducing variation. $F = (5545)/(130) = 42.6$, with 1 and 35 d.f., shows that b is highly significant.

The analysis thus far shows clearly that the drugs produced important effects on the denervated muscles and that these effects were all harmful, in the sense that they hastened loss of weight. Reference to Table I shows that this loss of weight was, to some extent, general, because total body weight and the intact muscle weight both decreased under some of the treatments. It is reasonable to ask, then, if the drugs had any specific effect on the denervated muscle, because it may be that the excessive weight loss of the denervated muscle was simply the result of the decrease in total body weight.

This question may be explored by following a little further the approach which has already been employed, that is, by fitting a regression equation to express the dependence of final denervated muscle weight on both initial and final total body weights and examining the residual variation. Final total body weight (y) was recorded in the experiment with this purpose in view.

The regression equation may be written

$$U - \bar{u} = b_1(x - \bar{x}) + b_2(y - \bar{y}).$$

The coefficients b_1 and b_2 are determined by the normal equations

$$(x^2)b_1 + (xy)b_2 = (xu)$$

$$(xy)b_1 + (y^2)b_2 = (yu)$$

in which the bracketed symbols denote sums of squares and products of deviations from means. The sum of squares of residuals is given by $(u^2) - b_1(xu) - b_2(yu)$, which becomes, on substituting for b_1 and b_2,

$$(u^2) - \frac{(x^2)(yu)^2 - 2(xy)(xu)(yu) + (y^2)(xu)^2}{(x^2)(y^2) - (xy)^2}.$$

All this is standard regression theory. This formula may be used to partition the total s.s. (u^2) into two parts, one attributable to regression and the other representing deviations from regression. The arithmetic is carried out exactly according to the pattern of Tables IV and V. The results are given in Tables VII and VIII.

TABLE VII

SUMS OF SQUARES AND PRODUCTS

	d.f.	(x^2)	(xy)	(y^2)	(xu)	(yu)	(u^2)
times	2	264	287	414	−1736	−1450	13269
drugs	3	2969	8543	35228	3451	15782	7931
times × drugs	6	9053	6671	8087	2083	2457	1529
error	36	38244	26612	25108	14562	11972	10111
times + error	38	38508	26899	25522	12826	10522	23380
drugs + error	39	41213	35155	60336	18013	27754	18042
times × drugs + error	42	47297	33283	33195	16645	14429	11640

TABLE VIII

SUMS OF SQUARES FOR REGRESSION AND DEVIATIONS FROM REGRESSION

row		regression		deviations		
		d.f.	s.s.	d.f.	s.s.	m.s.
1	error	2	6058	34	4053	119
2	times + error	2	4635	36	18745	
3	drugs + error	2	12930	37	5112	
4	times × drugs + error	2	6612	40	5028	
5	times			2	14692	7346
6	drugs			3	1059	353
7	times × drugs			6	975	162

It is interesting to observe the effect of introducing the final total weight into the analysis. The error mean square has not been affected appreciably (130 to 119), but the drugs mean square has been reduced from 1868 to 353. The ratio $F = (353)/(119) = 2.97$, with 3 and 34 d.f., is just below the 5% point, but whether or not this is judged to be significant, it is clear that by far the greater part of the differences among drugs is accounted for by their effects on the total weights of the animals. Values of denervated muscle weight, adjusted to equalize both initial and final total body weights, are given in Table IX.

We have, in the foregoing analysis, examples of the two chief uses of the analysis of covariance. The initial total weights were brought into the analysis to provide control over a source of disturbance which did not lend itself to experimental control. It is characteristic of variables

TABLE IX
DENERVATED MUSCLE WEIGHTS, ADJUSTED FOR INITIAL AND FINAL TOTAL BODY WEIGHTS

	A	B	C	D
4 days	429	448	435	526
8 days	311	368	360	586
12 days	241	231	323	557

which are used for this purpose that they are independent of treatment effects. Here the interpretation of the analysis is clear-cut. We are not, as a rule, concerned with testing the dependence of one variable on another. Rather, we require *a priori* confidence that the variable selected *does* furnish the desired control.

The final total weights were introduced for a wholly different purpose, not to establish control over a source of variation, but to aid in understanding the results of the experiment. Here, the earlier analysis established the fact that the drugs did produce substantial effects on the weights of the denervated muscles. The later analysis was undertaken to find out the manner in which the effects were brought about. The variable introduced for this purpose, final total weight, is affected by the treatments. This is more or less typical in this second use of covariance analysis. In cases of this kind, considerable caution must be used, both in deciding whether or not to introduce the variable and in the interpretation of the results.

We have not, in this second case, calculated a significance test either for b_2 or for the combined effects of b_1 and b_2. These tests are easily computed, but there is no reason to do so here, since none of our conclusions depends on tests of these quantities. This need not always be the case. Indeed, in some instances, the primary reason for conducting an analysis of covariance is to perform a test of significance. A case of this kind occurs in the experiment we have been discussing.

Shortly after denervation, a muscle exhibits a random twitching called fibrillation, which persists with gradually decreasing intensity until atrophy is complete. It has been conjectured that fibrillation causes atrophy and therefore, that a treatment which diminishes the intensity of this twitching should delay the progress of atrophy. The drugs used in this study were chosen because of their known effects on fibrillation. Measurements of the intensity of fibrillation (w) were made by electrical methods. The records are given in Table I.

If a lowering of the intensity of fibrillation does in fact decrease the rate of atrophy, we should find a negative correlation between the variables u and w. It is clear, however, that a correlation coefficient calculated in the ordinary way from the 48 pairs of observations on u and w may be misleading, because each of the variables in the correlation is affected both by the drugs and by the elapsed time. These sources of disturbance may be avoided by using the sums of squares and products from the error row to calculate the correlation coefficient. (Of course, a regression coefficient could be used here instead of a correlation coefficient.) This computation, then, will be based on Table X.

TABLE X

SUMS OF SQUARES AND PRODUCTS

	d.f.	(u^2)	(uw)	(w^2)
times	2	13269	801.8	67.28
drugs	3	7931	−176.6	16.81
times × drugs	6	1529	182.2	310.55
error	36	10111	−170.1	520.19

The coefficient of correlation calculated from the error row is

$$r = \frac{-170.1}{[(10111)(520.19)]^{1/2}} = -.0742$$

which may be tested for significant departure from zero by entering a table prepared for this purpose with $n = 35$ (e.g. R. A. Fisher, Statistical Methods for Research Workers, Table V. A.). This value of r is much too small to be judged significant.

Only the last row of Table X has been used in this calculation. Ordinarily, therefore, these numbers would be calculated directly, omitting the other rows of Table X. However, when interaction terms supply the error row, it is simplest to calculate the whole table. Also, it is occasionally of interest to calculate a correlation or regression coefficient from rows other than the error row. In Table X, the times row yields a correlation coefficient $r = .8478$ and in the drugs row, $r = -.4737$. These correlations are based on too few degrees of freedom to warrant much discussion, but their magnitudes show the importance of removing these degrees of freedom before attempting to test for relationship between u and w. The correlation calculated from the error row is a *partial* correlation in which the effects of time and drugs have been eliminated.

Doubtless it is desirable that we remove also the effects of the dependence of u and w on the variable x (initial total body weight). This amounts to fitting regressions of u and v on x and correlating the residuals. The computation may be regarded as an analysis of covariance according to the pattern laid out in Tables XI and XII.

TABLE XI
SUMS OF SQUARES AND PRODUCTS

	d.f.	(u^2)	(uw)	(w^2)	(ux)	(wx)	(x^2)
times	2	13269	801.8	67.28	−1736	−72.4	264
drugs	3	7931	−176.6	16.81	3451	−195.8	2969
times × drugs	6	1529	182.2	310.55	2083	266.7	9053
error	36	10111	−170.1	520.19	14562	200.0	38244

TABLE XII
SUMS OF SQUARES AND PRODUCTS FOR REGRESSIONS AND DEVIATIONS FROM REGRESSIONS

	regression				deviations			
	d.f.	(u^2)	(uw)	(w^2)	d.f.	(u^2)	(uw)	(w^2)
error	1	5545	76.2	1.05	35	4566	−246.3	519.14

The correlation coefficient calculated from the residuals is then

$$r = \frac{-246.3}{[(4566)(519.14)]^{1/2}} = -.1600$$

This number would be tested by entering the table with $n = 34$. It is, of course, far from attaining significance and therefore we find no convincing evidence of a connection between atrophy and fibrillation.

The only calculation in Table XII which has not been encountered earlier is that which yields the sum of products of deviations about regressions. This number is obtained by combining sums of squares and products taken from the error row of Table XI.

$$-170.1 - (14562)\,(200.0)/(38244) = -170.1 - 76.2 = -246.3$$

These examples illustrate the common types of analysis in which covariance methods are used. Occasionally more elaborate applications are required. For example, it may happen that a linear regression is incapable of expressing adequately the dependence of one variable on another and therefore a curved regression line is required. Procedures appropriate to these unusual applications may be derived without difficulty from the theory of least squares, but no attempt will be made here to discuss questions of this kind.

SECTION II

The analysis is carried, in Section I, only to the point where standard covariance procedures are illustrated. The results obtained there indicate that further examination of the data is called for. Table IX suggests that not all the differences among drugs are accounted for by variation in initial and final body weights. Furthermore, it may be that final body weight is not the best variable to use in making allowance for general drug effects, since it is to be expected that final denervated muscle weight would be more highly correlated with final intact muscle weight than with final body weight. At any rate, it should be worth while to repeat the computations which produced Tables VIII and IX, using final intact muscle weight in place of final body weight. The results are shown in Tables XIII and XIV.

TABLE XIII
SUMS OF SQUARES OF DEVIATIONS FROM REGRESSION

	d.f.	s.s.	m.s.
error	34	2483	73
times	2	10132	5066
drugs	3	480	160
times × drugs	6	939	156

TABLE XIV
DENERVATED MUSCLE WEIGHTS, ADJUSTED FOR INITIAL BODY WEIGHT AND FINAL INTACT MUSCLE WEIGHT

	A	B	C	D
4 days	430	441	440	444
8 days	360	347	374	312
12 days	327	246	340	246

The error term in Table XIII confirms the guess that intact muscle weight is more closely correlated with denervated muscle weight than is final body weight, since the residual error mean square is reduced to 73, in contrast with 119 of Table VIII. The "drugs" mean square is seen to be not significant with an F value of 2.19. However, Table XIV indicates that, when allowance is made for general drug effects, using intact muscle weight, the two levels of atropine (A and C) have practically the same effect, that the effect of quinidine (B) differs little from that of saline (D) and that the effect of atropine is appreciably different from those of quinidine and saline. This difference is so marked as to suggest the possibility that one of the three "drugs" degrees of freedom may be reflecting a real difference and ought to be examined by itself.

Numerical tests support the decisions reached by inspection of Table XIV and indicate also that the average time curve for u, adjusted for the values of x and v, is practically linear. It seems reasonable, therefore, to direct attention to 3 of the 11 "treatment" degrees of freedom, the atropine-saline contrast, symbolized by $A + C - 2D$, the linear component of the time curve, $t_3 - t_1$, and their interaction, $(A + C - 2D)(t_3 - t_1)$. The corresponding mean squares, adjusted by covariance for variation in the values of x and v, are 183, 9863 and 508. Comparison of these mean squares with the error mean square (73) yields the F-values 2.51, 135.03, 6.95, each with 1 and 34 d.f. The component $A + C - 2D$ does not attain significance, but its interaction with linear time is well above the 5% point. The reason for this is clear. Differences between responses to atropine and to saline are small initially and increase with time to the point where they become fairly large by the end of the twelfth day. These final large differences are masked somewhat in the main effect, since they are considered only in conjunction with the smaller differences of the earlier periods, but emerge sharply in the interaction component where they are *contrasted* with the differences at the 4 day period. It appears, therefore, that this interaction component is the critical element in the analysis and we are led to the decision that, when allowance is made for general weight loss caused by the drug, the weight of the denervated muscle decreases somewhat more slowly when atropine is used than in the controls. At this point, we find ourselves offering the same conclusions as were reached on the basis of Table II and we may well be disturbed here, as we were before, at finding that those muscle weights which diminished most turn out, after adjustment, to show the smallest decrease. Certainly, when we give the verdict "A and C are alike and B and D are alike, but A and C differ from B and D", we are making a statement which obviously is true for the *intact* muscles, since under treatments A and C, intact muscle weight decreases, whereas under treatments B and D it shows a steady increase.

Could it be this fact alone which accounts for the differences among the u-values in Table XIV? If so, our method of adjustment must be at fault. In any event, it becomes necessary to inquire into the assumptions on which the covariance adjustments rest and to find out if they are satisfied in this experiment. Section III presents an attempt to test some of these assumptions and to search into some of the physiological implications of the results of the experiment.

SECTION III

Naturally, a covariance adjustment is likely to be misleading unless the form of the function used to depict the relationship of the dependent to the independent variables (in these examples a linear function) is such as to represent this relationship adequately. Usually, with experimental data, a linear function can be made to serve reasonably well, by taking care, in the planning of the experiment, to avoid too large a range in the independent variables. Tests of linearity may be made exactly as in ordinary regression analysis and will not be discussed here.

Another difficulty, more troublesome to deal with, arises from the fact that, even though the regressions are linear, the regression coefficients may vary from one part of the experiment to another. For example, the dependence of u on x and v may vary with time or from one drug to another. Indeed, it seems not unlikely that variation of this kind would be encountered in this experiment, but the covariance adjustments which have been made are based on the assumption that a single set of regression coefficients can be used throughout the experiment.

A test for homogeneity of regression coefficients can be made by standard covariance methods [2]. These tests show that the regression coefficients are substantially the same for all four drugs, but that they tend to decrease in value with increasing time. It seems appropriate, therefore, to use different regressions for the three times. When this is done, the numbers in Table XIV are not altered appreciably, so that we may accept the findings based on Tables XIII and XIV and attempt an interpretation of them.

An inquiry into the physiology of this situation reveals that little meaning can be attached to values of denervated muscle weight adjusted for either final intact muscle weight or final body weight, because the weight loss of the denervated muscle is composed of two portions, that attributable to atrophy and that caused by the drug (or, negatively, by growth). Weight loss due to atrophy is permanent, under the conditions of this experiment, whereas weight changes brought about by the drugs or by growth are recoverable. Furthermore, after a muscle fibre has atrophied, its weight is almost entirely unaffected by the administration of drugs or by growth of the animal. Thus, the denervated muscle becomes less responsive to changes from these sources as atrophy progresses. Doubtless this fact accounts for the contrast between drugs A and C and drugs B and D in Table XIV. It appears, then, that for an investigation of atrophy, we have been studying the wrong variable (denervated muscle weight) and using the wrong statistical model (linear regression with constant coefficients of denervated muscle weight on initial body weight and final body weight or final intact muscle weight). It is clear, too, that we cannot undertake an analysis which will bring out the behaviour of atrophy until we have an appropriate model on which to base the analysis.

Apparently not enough quantitative information is available to support any specific model. The following speculations are offered simply to bring out the nature of the problem.

Let us suppose that a muscle fibre which atrophies, decreases in weight in the ratio of $\lambda : 1$ and that the ultimate weight of the fibre is not affected by the drugs or by growth. Let p_t represent the *proportion* of the muscle fibres which have atrophied after a period of denervation of duration t. Then this portion of the muscle weighs $\lambda u_0 p_t$ where u_0 is the initial weight of this muscle. The remaining portion of the muscle, which has not atrophied, weighs $(1 - p_t)u_0 f(t)$, where the function $f(t)$ is introduced to take account of changes in weight due to drugs, growth and possibly other effects also. (No doubt even the functional form of $f(t)$ changes from one treatment to another.) Then the weight of the denervated muscle at time t is given by

$$u_t = \lambda u_0 p_t + u_0(1 - p_t)f(t).$$

Similarly, the weight of the intact muscle, at time t, is

$$v_t = v_0 f(t).$$

Now, if we equate u_0 and v_0 (which should introduce no serious error) and eliminate $f(t)$ between the two equations, we obtain

$$u_t = \lambda u_0 p_t + (1 - p_t)v_t .$$

The initial weight of the denervated muscle, u_0 , is not measurable and its value varies from one animal to another, but it seems reasonable to assume, as we did earlier, that initial denervated muscle weights vary linearly with initial body weights, that is,

$$u_0 = a + bx.$$

Also, it seems worth while to try the assumption that p_t increases linearly with time, since the time range in this experiment is relatively short. Thus, writing $p_t = \alpha t$ and substituting for p_t and u_0, the equation may be written

$$\frac{v_t - u_t}{t} = -\lambda a\alpha - \lambda b\alpha x + \alpha v_t \,.$$

Now the assertion that the various drugs produce different effects on the progress of atrophy is equivalent to the statement that p_t varies from one drug to another. On the other hand, if p_t is unaffected by the drugs and if all the assumptions which have been made are valid, than α is constant throughout the experiment and a single regression equation of $(v_t - u_t)/t$ on x and v_t should fit all of the observations reasonably well. This may be tested by applying to $z = (v_t - u_t)/t$ the same covariance analysis as was conducted on u_t in Section II. The results of the calculations are given in Tables XV and XVI. (The t values used in computing the z's are 1, 2, 3.)

TABLE XV

SUMS OF SQUARES OF DEVIATIONS FROM REGRESSION

	d.f.	s.s.	m.s.
error	34	933	27
times	2	168	84
drugs	3	113	38
times × drugs	6	69	11

TABLE XVI

VALUES OF z, UNADJUSTED AND ADJUSTED FOR INITIAL BODY WEIGHT AND FINAL INTACT MUSCLE WEIGHT

	A		B		C		D	
	z	z'	z	z'	z	z'	z	z'
4 days	11.75	13.38	14.75	13.13	9.75	11.00	19.25	15.05
8 days	9.00	16.41	21.00	17.27	12.75	14.01	28.00	21.84
12 days	7.50	16.59	20.50	20.51	10.50	13.37	25.25	17.45

The "drugs" mean square is far from attaining significance; the "times" mean square is just below the 5 per cent point, which, even though not significant, may indicate that the assumption that $p_t = \alpha t$ is not too well satisfied. While this test cannot be considered a critical test of the effects of the drugs, at least we find nothing to support the view that the drugs have had a specific effect on the rate of atrophy. It seems likely that a much more elaborate and carefully designed experiment is needed to produce satisfactory evidence on this question.

Probably some of the discussion which led to this statistical model runs counter to facts which are known to physiologists, but this does not necessarily imply that the model is wholly incorrect. Alternative statements, more in accord with known facts but of a more complicated nature, can lead to the same model. A few specific points, however, deserve some comment.

The statement that, when a *muscle fibre* atrophies, its weight decreases in the constant ratio $\lambda : 1$ is an unsupported assumption, but at any rate, it seems less serious than the same assumption applied to the whole muscle (which was rejected early in this discussion), in view of the fact that the whole muscle contains, in addition to muscle fibres, material which does not lose weight as a result of denervation. This same point comes in again later on, when the weights of the muscles are treated as if they are made up entirely of the weights of muscle fibres. The inaccuracy introduced into the model by neglecting non-muscular elements may not appear in the final computations. At least, it is unlikely to affect the difference $v_t - u_t$, and its effects elsewhere may be taken up by the regression on initial weights. At any rate, this experiment provides no information with which this question can be investigated.

REFERENCES

[1] Solandt, D. Y., D. B. DeLury and John Hunter, "The Effect of Atropine and Quinidine Sulphate on Atrophy and Fibrillation in Denervated Skeletal Muscle"—*The American Journal of Physiology*, Vol. 140, No. 2, 1943. (A study conducted under a grant from the National Foundation for Infantile Paralysis, Inc.)

[2] Goulden, C. H., *Methods of Statistical Analysis*, New York, John Wiley and Sons, 1939. Chap. XV, Sec. 5.

AN EXPERIMENTAL DESIGN USED TO ESTIMATE THE OPTIMUM PLANTING DATE FOR COTTON

R. L. Anderson

Institute of Statistics, Raleigh, N. C.

H. L. Manning

Empire Cotton Growing Corporation and Agricultural Department
Uganda, British East Africa

INTRODUCTION

Two experiments were conducted in 1946–1947 by the Agricultural Research Station of the Empire Cotton Growing Corporation in the Uganda Protectorate of British East Africa to determine the optimum planting date for cotton. One experiment was laid out at the Kawanda Station, where the main rains occur during the first half of the year; the other experiment was laid out at Kabula, where the main rains occur during the second half of the year. Since it was suspected that interference from the insect, *Lygus simonyi*, would be a potent factor, it was not considered advisable to plant a long series of sowing dates in one locality. Lygus is a small capsid bug, which breeds upon grain crops planted with the first rains in a given locality; when these early crops are harvested, the bug migrates on to newly planted cotton, if such is available. There has also been some scanty evidence that where young cotton is planted in contiguous plots, the Lygus bugs would tend to migrate to the younger cotton.

In these experiments, it was decided to plant at three successive dates, separated by intervals of two weeks (one fortnight), at each of eight localities, with a two-thirds overlap between successive localities. That is, dates 1, 2 and 3 were to be used at the first locality; dates 2, 3 and 4 at the second locality; dates 3, 4 and 5 at the third locality; etc. This experimental design allowed us to use ten planting dates, spread over an 18 week period. Four replications were used at each locality, with the three dates at each locality being used in each replicate. Hence the yields at each locality could be analyzed on the basis of a 4 × 3 randomized blocks design. The design and actual plot yields (kilograms of seed cotton per plot) of the experiment at the Kawanda Station are presented in Table 1. The same design was used at the Kabula Station, except that the planting dates ranged from September 13 through January 17.

TABLE 1

DESIGN AND ACTUAL YIELDS (KGMS/PLOT) OF SEED COTTON FOR A PLANTING DATE TRIAL AT THE KAWANDA STATION*

	Dates										
Locality	1 May 1	2 May 15	3 May 29	4 June 12	5 June 26	6 July 10	7 July 24	8 Aug. 7	9 Aug. 21	10 Sept. 4	Total
1	3.35	3.86	1.99								
	1.49	2.71	2.89								
	2.44	2.18	1.68								
	2.44	1.95	2.13								
	9.72	10.70	8.69								29.11
2		8.84	7.02	7.02							
		6.23	7.93	7.69							
		7.05	6.60	6.61							
		6.55	6.12	5.19							
		28.67	27.67	26.51							82.85
3			6.12	6.34	3.83						
			5.67	5.67	4.30						
			5.22	5.66	5.20						
			6.10	5.17	4.33						
			23.11	22.84	17.66						63.61
4				2.45	3.17	1.94					
				2.80	1.50	1.39					
				2.47	1.95	1.54					
				2.71	1.96	3.19					
				10.43	8.58	8.06					27.07
5					3.85	4.94	3.85				
					4.57	3.53	3.21				
					4.56	4.53	3.82				
					4.70	5.27	3.72				
					17.68	18.27	14.60				50.55
6						5.21	3.75	3.25			
						4.31	4.62	3.00			
						3.81	3.40	3.38			
						4.14	4.17	2.96			
						17.47	15.94	12.59			46.00
7							5.66	5.17	4.94		
							6.06	4.17	3.58		
							5.66	5.94	4.26		
							7.01	5.58	4.62		
							24.39	20.86	17.40		62.65
8								5.11	3.84	2.76	
								4.51	3.55	3.05	
								4.75	3.93	3.39	
								4.18	3.70	2.99	
								18.55	15.02	12.19	45.76
Total	9.72	39.37	59.47	59.78	43.92	43.80	54.93	52.00	32.42	12.19	407.60

*The plot size was 12′ x 42′, approximately 1/86 acre.

The problem is to determine the best planting date, and to set up some kind of confidence limits on this date. One also might want to consider the relative efficiency of this design relative to other designs, assuming that the insect interference were actually negligible. For example, what could one gain if he planted at four successive dates at each locality, giving a three-fourth overlap? Also what is the loss in efficiency as compared to a completely balanced design, whereby all dates are used at each locality, assuming no insect interference?

ANALYSIS OF THE KAWANDA DATA

(a) **Analysis of Variance.** In order to evaluate the true differences between the yields at successive planting dates, it is necessary to adjust the average yields for the differences among the localities. This necessitates a rather complicated least squares solution of the date and locality effects, which we could avoid if there was good evidence of the non-existence of real locality differences. As a preliminary step, we shall consider the analysis of variance for the Kawanda data. It will be granted that we have not demonstrated the non-existence of locality effects even though the analysis of variance shows no significant differences; however, the experimentalist probably would be willing to neglect small locality effects in order to forego the necessity of carrying out the least squares solution.

From Table 1 we note that there are 23 degrees of freedom for the locality, date and residual (locality-date interaction) constants. Since there are 7 degrees of freedom for localities and 9 for dates, there must be 7 remaining for the locality-date interaction. The sum of squares for these 7 degrees of freedom can be determined from the sum of 7 independent squares, each representing one degree of freedom. These 7 independent squares are formed by squaring various linear combinations of the 24 locality-date totals. Let y_{tj} represent the total yield of the 4 plots planted at the t-th date and the j-th locality, where $t = 1, 2, \cdots, 10$ and $j = 1, 2, \cdots, 8$. A linear combination C can be represented as

$$C = \sum a_{tj} y_{tj} . \tag{1}$$

Then $C^2/4 \sum a_{tj}^2$ is a quantity which can be used to represent the sum of squares with one degree of freedom. In forming the 7 independent linear combinations, 7 different sets of a_{tj} are needed. In order that any two forms C and C' be independent,

$$\sum a_{tj} a'_{tj} = 0.$$

TABLE 2

THE COEFFICIENTS TO FORM A SET OF 7 INDEPENDENT LINEAR COMBINATIONS FOR THE LOCALITY-DATE INTERACTIONS

Linear Combinations*

Locality	Date	Order of Planting	Yield	1	2	3	4	5	6	7
1	2	2	10.70	+1				+1		+1
	3	3	8.69	−1				−1		−1
2	2	1	28.67	−1				−1		−1
	3	2	27.67	+1				−3		−3
	4	3	26.51					4		4
3	3	1	23.11					4		4
	4	2	22.84		+1			−3		11
	5	3	17.66		−1			−1		−15
4	4	1	10.43		−1			−1		−15
	5	2	8.58		+1			1		−41
	6	3	8.06							56
5	5	1	17.68							56
	6	2	18.27			+1			1	−41
	7	3	14.60			−1			−1	−15
6	6	1	17.47			−1			−1	−15
	7	2	15.94			+1			−3	11
	8	3	12.59						4	4
7	7	1	21.39						4	4
	8	2	20.86				+1		−3	−3
	9	3	17.40				−1		−1	−1
8	8	1	18.55				−1		−1	−1
	9	2	15.02				+1		+1	+1
Total (C)				1.01	3.33	2.14	−0.07	0.78	2.79	17.99
Divisor (D)				16	16	16	16	224	224	43,456
C^2/D				1.0433				0.0375		0.0074

*Blank spaces represent 0 coefficients. Each yield is the total yield of 4 plots at the Kawanda Station.

In addition, these combinations must be independent of date and locality effects. In order to fulfill this condition, the a_{ij} of a given C for any date or locality must also sum to 0.

One set of 7 independent linear combinations for the residual or interaction sum of squares is given in Table 2. The first four combinations represent first-order interaction effects. For example consider

$$(2) \qquad C_1 = (y_{21} - y_{31}) - (y_{22} - y_{32}) = y_{21} - y_{31} - y_{22} + y_{32} \,.$$

Each part of the right-hand side measures the difference between the yields at dates 2 and 3 but at two different locations, 1 and 2. Hence the difference between these two parts measures the change in the date effect from one location to another, which we designate as the date-location interaction. We note that C_1 is independent of date and locality effects, since the sum of the coefficients for each date and locality is zero. Also the first four combinations are obviously independent, since they have no plots in common.

TABLE 3

ANALYSIS OF VARIANCE FOR THE KAWANDA DATA

Source	Degrees of Freedom	Sum of Squares	Mean Square
Interaction	7	1.0882	.1555
Localities	7	201.1345	
Dates (adj.)	9	21.7354	2.4150**
Dates	9	45.2186	
Localities (adj.)	7	177.6513	25.3787**
Replications	24	11.6219	0.4842
Error	48	17.3125	0.3606

**Significant at the 1% probability level.

In this experiment these first four combinations also represent some effect of the order of planting on yield (and consequently of Lygus migration from the early to the late plantings). The effect of the Lygus migration from the early to the late plantings is a quadratic effect, since the sums (C) are sums of single yields at the first and third plantings subtracted from two yields at the second planting. The sum of the first four combinations could be used to measure this quadratic effect of the order of planting with one degree of freedom, and the remaining three degrees of freedom would then represent the interaction effect, adjusted for order of planting. The sum of squares for the order of planting would be $(6.41)^2/4(16) = .6420$, leaving 0.4013 for the three interaction degrees of freedom. None of these effects is significantly different from zero, using the error variance given in Table 3. Since there is obviously no date-location interaction, we can use each of the first four C's as a measure of the effect of order of planting. The standard error of each of these C's is 2.40, indicating that none of them is significantly different from zero. It is possible to compute a separate error term for each C with 6 degrees of freedom, but this does not alter any of the above conclusions.

Combinations 5 and 6 are each based on a comparison of 4 dates and 4 localities. For example

$$C_5 = y_{21} - y_{22} - y_{31} - 3y_{32} + 4y_{33} \tag{3}$$
$$+ 4y_{42} - 3y_{43} - y_{44} - y_{53} + y_{54} \,.$$

Again the sum of the coefficients for any date or locality is 0. In addition the sum of the products of these coefficients with those for any of the first 4 combinations is 0. Combination 7 is an over-all comparison, which was constructed so as to be independent of all the other combinations as well as dates and localities.

The divisors are the denominators, $4\sum a_{ij}^2$, mentioned above. For the first 4 combinations the divisors are $4(1 + 1 + 1 + 1) = 16$. For the next two, the divisors are $2[4(1 + 1 + 1 + 9 + 16)] = 224$. And finally for the last comparison, the divisor is $4(10864) = 43456$. The combination totals and contributions to the residual sum of squares are given at the bottom of Table 2, and $SSR = 1.0882$.

Next we compute the total sum of squares for the 23 degrees of freedom (SST) and the sum of squares for the unadjusted date effects (SSD), and the sum of squares for the unadjusted locality effects (SSL). These are computed as follows:

$$\text{Correction for mean} = C = (407.60)^2/96 = 1730.6017$$

$$SST = [(9.72)^2 + (10.70)^2 + \cdots + (12.19)^2]/4 - C$$
$$= 223.9581$$

$$SSL = [(29.11)^2 + (82.85)^2 + \cdots + (45.76)^2]/12 - C \tag{4}$$
$$= 201.1345$$

$$SSD = \frac{(9.72)^2}{4} + \frac{(39.33)^2}{8} + \frac{(59.47)^2}{12} + \cdots + \frac{(12.19)^2}{4}$$
$$- C = 45.2186$$

In order to make exact tests of significance, we require the sum of squares for localities adjusted for dates [SSL(adj.)] and the sum of squares for dates adjusted for localities [SSD(adj.)]. A simple identity can be used to evaluate these adjusted sums of squares:

$$\begin{aligned} SSL + SSD(\text{adj.}) &= SSD + SSL(\text{adj.}) \\ &= SST - SSR = 222.8699 \end{aligned} \tag{5}$$

Hence

$$\begin{aligned} SSD(\text{adj.}) &= 222.8699 - 201.1345 = 21.7354 \\ SSL(\text{adj.}) &= 222.8699 - 45.2186 = 177.6513. \end{aligned} \tag{6}$$

Finally the error sum of squares (SSE) is given by the replication by date sum of squares summed over the 8 locations, with $3 \times 2 \times 8 = 48$ degrees of freedom. Because of the large locality differences, it might be suspected that the error variance would not be the same for all locations. Bartlett's χ^2 test for the homogeneity of variance was applied, giving $\chi^2 = 7.364$ with 7 degrees of freedom. The probability of obtaining this or a larger value of χ^2, assuming equal error variances at all localities, is about .40, indicating that it is quite reasonable to assume equal error variances at all localities and to pool the individual estimates as mentioned above. It might be argued that we should also pool the location × date interaction with the error variance, giving a total of 55 error degrees of freedom. We have not done this in the analysis which follows, because the additional degrees of freedom were not thought necessary in this case.

The complete analysis of variance is given in Table 3.

This analysis indicates that there are highly significant differences among both the dates and localities. Since there are real differences between the mean yields for different localities, the mean yields at each planting date should be adjusted for locality effects. Two types of analysis are suggested at this point. Either we determine the adjusted average yield for each planting date, or we fit a regression curve of some kind to the adjusted average yields.

(b) Average Yields for Each Planting Date, Adjusted for Locality Effects. We can represent the total yield y_{tj} of the 4 plots at a given date, t, and locality, j, as follows:

$$y_{tj} = 4(m + d_t + l_j) + r_{tj} , \tag{7}$$

where m is the general mean, d_t is the effect of the t-th planting date ($t = 1, 2, \cdots, 10$) and l_j is the effect of the j-th locality ($j = 1, 2, \cdots, 8$), r_{tj} is the residual after accounting for the effects of the constants—m, d_t, and l_j. It should be noted that we have multiplied the right-hand side by 4 in order to put the results on a per-plot basis.

The constants will be estimated by the method of least squares. First we set up the error equation

$$SSR = \sum [y_{ti} - 4(m + d_t + l_j)]^2/4 \tag{8}$$

where SSR indicates the interaction or residual sum of squares, the sum of the 24 squared residuals, $r_{tj}^2/4$. The constants are estimated by minimizing SSR with respect to each of the constants. For d_1, we set

$$\frac{\partial SSR}{\partial d_1} = 0. \tag{9}$$

This gives

$$\sum_j [y_{1j} - 4(m + d_1 + l_j)] = 0, \tag{10}$$

where the summation is made over only the one locality (4 plots) having the first planting date. Equation (10) simplifies to

$$\sum_j y_{1j} = D_1 = 4d_1 + 4l_1 + 4m \tag{11}$$

Similar equations can be obtained for each d_t.

Similarly if we minimize SSR with respect to l_1, we have

$$\sum_t y_{t1} = L_1 = 4(d_1 + d_2 + d_3) + 12l_1 + 12m. \tag{12}$$

The equation for m is

$$\begin{aligned}\sum_{tj} y_{tj} = G = 4d_1 + 8d_2 + 12(d_3 + \cdots + d_8) + 8d_9 + 4d_{10} \\ + 12(l_1 + \cdots + l_8) + 96m\end{aligned} \tag{13}$$

The coefficients of all 19 least squares equations are presented in Table 4. The reduced equations, after eliminating the locality constants, are given at the bottom of Table 4. In this table the yield totals, adjusted for localities, are indicated by A_t. The G equation indicates that if m is to be the general mean, we should assume that $\sum_j l_j = 0$ and that

$$[d_1 + d_{10} + 2(d_2 + d_9) + 3(d_3 + \cdots + d_8)] = 0. \tag{14}$$

In order to obtain standard errors to test for the difference between the adjusted average yields at any two dates, it is necessary to invert the coefficient matrix for the adjusted date effects. The inverted matrix is presented in the upper part of Table 5. The elements of this matrix (which are indicated by the symbol m_{tj}, where t stands for the row and

TABLE 4

COEFFICIENTS OF LEAST SQUARES EQUATIONS FOR ESTIMATION OF PLANTING DATE MEANS

Independent Variables

Plot Totals*	Dates (d_i) 1	2	3	4	5	6	7	8	9	10	Localities (l_j) 1	2	3	4	5	6	7	8	m
D_1 : 9.72	4										4								4
D_2 : 39.37		8									4	4							8
D_3 : 59.47			12								4	4	4						12
D_4 : 59.78				12								4	4	4					12
D_5 : 43.92					12								4	4	4				12
D_6 : 43.80						12								4	4	4			12
D_7 : 54.93							12								4	4	4		12
D_8 : 52.00								12								4	4	4	12
D_9 : 32.42									8								4	4	8
D_{10} : 12.19										4								4	4
L_1 : 29.11	4	4	4								12								12
L_2 : 82.85		4	4	4								12							12
L_3 : 63.61			4	4	4								12						12
L_4 : 27.07				4	4	4								12					12
L_5 : 50.55					4	4	4								12				12
L_6 : 46.00						4	4	4								12			12
L_7 : 62.65							4	4	4								12		12
L_8 : 45.76								4	4	4								12	12
G : 407.60	4	8	12	12	12	12	12	12	8	4	12	12	12	12	12	12	12	12	96

EQUATIONS FOR DATE EFFECTS ADJUSTED FOR LOCALITIES

Adjusted Totals*		Coefficients of Adjusted Date Effects (d'_i) 1	2	3	4	5	6	7	8	9	10
$A_1 = 3D_1 - L_1$	= 0.05	8	−4	−4							
$A_2 = 3D_2 - L_1 - L_2$	= 6.15	−4	16	−8	−4						
$A_3 = 3D_3 - L_1 - L_2 - L_3$	= 2.84	−4	−8	24	−8	−4					
$A_4 = 3D_4 - L_2 - L_3 - L_4$	= 5.81		−4	−8	24	−8	−4				
$A_5 = 3D_5 - L_3 - L_4 - L_5$	= −9.47			−4	−8	24	−8	−4			
$A_6 = 3D_6 - L_4 - L_5 - L_6$	= 7.78				−4	−8	24	−8	−4		
$A_7 = 3D_7 - L_5 - L_6 - L_7$	= 5.59					−4	−8	24	−8	−4	
$A_8 = 3D_8 - L_6 - L_7 - L_8$	= 1.59						−4	−8	24	−8	−4
$A_9 = 3D_9 - L_7 - L_8$	= −11.15							−4	−8	16	−4
$A_{10} = 3D_{10} - L_8$	= −9.19								−4	−4	8

*Plot Totals at Kawanda.

j the column number) can be used to determine both the values of the adjusted date effects and their variances. Before inverting the 10-row coefficient matrix at the bottom of Table 4, it is noted that the coefficients always sum to 0 for a given row or column. Such a matrix is

TABLE 5

INVERSE (M_{tj}) MATRIX FOR ADJUSTED DATE EFFECTS*

j

	1	2	3	4	5	6	7	8	9	10
1	.48884									
2	.37500	.40550								
3	.35268	.34450	.36086							
4	.30580	.30799	.30362	.31892						
t 5	.26562	.26504	.26622	.26208	.27744					
6	.22322	.22337	.22306	.22414	.22015	.23503				
7	.18304	.18299	.18308	.18276	.18396	.17949	.19615			
8	.13616	.13617	.13615	.13621	.13601	.13675	.13398	.14434		
9	.11384	.11383	.11385	.11379	.11399	.11325	.11602	.10566	.14434	
10	0	0	0	0	0	0	0	0	0	
A_j**	0:05	6.15	2.84	5.81	−9.47	7.78	5.59	1.59	−11.15	−9.19
d'_t	4.300	4.478	4.110	3.856	2.981	3.022	2.387	1.598	0.699	0
(d_t+m)	5.691	5.869	5.501	5.247	4.372	4.413	3.778	2.989	2.090	1.391
$s^2(d'_t)$	0.5288	0.4387	0.3904	0.3450	0.3001	0.2543	0.2122	0.1561	0.1740	0
$s(d'_t)$	0.7272	0.6623	0.6248	0.5874	0.5478	0.5043	0.4607	0.3951	0.4171	0

*Since the matrix is symmetrical, only the lower left section is reproduced here.

**Data from the Kawanda experiment.

called a *singular matrix* and cannot be inverted. This difficulty could have been taken care of by dropping the last equation and eliminating d_{10} by use of equation (14). However, we decided on a simpler procedure, namely to assume $d_{10} = 0$, and hence drop the 10-th row and column from the matrix. Then we merely inserted a column and a row of zeros in the inverted matrix. We have indicated the date effects under this assumption as d'_t .

For those not accustomed to the matrix notation, we might add that a matrix is simply a two-way array of figures such as the coefficients in Table 4. We are here dealing with symmetrical matrices. An inverted symmetrical matrix consists of another two-way array which has the following property: If you multiply the elements of any row (or column) of the original symmetrical matrix by the corresponding elements of the same row (or column) of the inverted matrix, the sum of the products is unity; if you multiply the corresponding elements of different rows or columns, the sum of the products is zero.

The values of the adjusted date effects are then computed from the M and A values as follows:

$$d'_t = \sum_{j=1}^{10} M_{tj} A_j \tag{15}$$

For example, $d'_1 = [(0.48884)(0.05) + \cdots + (0.11384)(-11.15)] = 4.3005$. Note that all $M_{t,10}$ values are 0. The complete set of d'_t values is presented in Table 5.

In order to obtain the original d_t values, we subtract a constant k from each d'_t so that the new values $d'_t - k = d_t$ fulfill equation (14). We see that

$$(16) \quad k = [d'_1 + 2d'_2 + 3(d'_3 + \cdots + d'_8) + 2d'_9 + d'_{10}]/24 = 2.855.$$

In most cases, it is desired to replace the date effects by adjusted yields at each date by adding m ($= 4.246$) to each of the d_t. Hence we can combine both steps by adding

$$(17) \qquad m - k = 4.246 - 2.855 = 1.391$$

to each of the d'_t. These adjusted yields for each date $(d_t + m)$ are also given in Table 5.

The variance of any adjusted date effect (d'_t) is simply $3M_{tt}\sigma^2$, where σ^2 is estimated by the error variance, 0.3606. The factor, 3, is required because the D_t were multiplied by 3 in determining the A_t. The variance of d'_1, for example, is

$$3(0.48884)(0.3606) = 0.5288$$

The variances and standard errors of d'_t are also given in Table 5.

However, we are generally more interested in the variances of the differences between the adjusted mean yields for successive dates. For example, the difference between the average yields for the first two dates is 0.18. The variance of the difference between the adjusted mean yields at two dates, i and k is given by the formula

$$(18) \qquad 3(M_{ii} - 2M_{ik} + M_{kk}]\sigma^2$$

where σ^2 is the error variance, estimated to be 0.3606. Hence the variance of the difference between the average yields for the first two dates is

$$3(0.48884 - 0.75000 + 0.40550)(0.3606) = (0.43301)(0.3606) = 0.1561.$$

Hence we conclude that there was no significant difference between the two adjusted mean yields. The coefficients of the error variance (for example, 0.43301) are given in Table 6.

Some of the differences between adjusted mean yields for successive dates and their standard errors are given in the second half of Table 6 on the next page.

TABLE 6

COEFFICIENTS OF VARIANCES OF DIFFERENCES BETWEEN ADJUSTED DATE MEANS

Date

	2	3	4	5	6	7	8	9	10
1	.43301	.43301	.58846	.70508	.83232	.95676	1.08256	1.21649	1.46652
2		.23205	.32532	.45860	.58137	.70701	.83250	.96651	1.21649
3			.21762	.31758	.44928	.57253	.69866	.83250	1.08256
Date 4				.21658	.31703	.44866	.57253	.70701	.95676
5					.21652	.31703	.44928	.58137	.83232
6						.21658	.31758	.45860	.70508
7							.21762	.32532	.58846
8								.23205	.43301
9									.43301

(19)

Dates	Mean Difference	Standard Error
1–2	−0.18	0.40
2–3	0.37	0.29
3–4	0.25	0.28
4–5	0.88	0.27
5–6	−0.04	0.27
6–7	0.63	0.27
7–8	0.79	0.28
8–9	0.90	0.29
9–10	0.70	0.40
4–6	0.84	0.34
5–7	0.59	0.34
2–4	0.62	0.34
2–5	1.50	0.41
2–6	1.46	0.46

From these averages, we conclude that: (i) The optimum planting date is somewhere near the first planting date (May 1). This optimum is poorly determined, as shown by the fact that there is no significant decrease in the adjusted yield until the fifth planting date (June 26). (ii) There is an unaccountable plateau at the fifth and sixth planting dates. Except for this, after the fourth date there appears to be a general significant decrease in yield between successive planting dates until we reach the last date (September 4). There is a tendency to flatten out at this last date, as expected, because the yields cannot fall below zero.

(c) Determination of the regression of cotton yields (adjusted for locality differences) on planting date. An examination of the adjusted mean yields $(d_t + m)$ in Table 5 suggests that a regression equation of the following type should be used:

$$20) \quad y_{ti} = 4[a + b(t - \bar{t}) + c(t - \bar{t})^2 + l_i] + p_{ti} ,$$

where $\bar{t}$ is the mean time period (= 5.5) and p_{tj} is the residual after accounting for the parabolic trend and the effect of the j-th locality, a, b and c are the coefficients of the regression curve, to be estimated from the data. As before, we shall assume that $\sum_j l_j = 0$.

The constants are estimated by least-squares. For example, the least-squares equation for b is

$$\sum [y_{tj}t' - 4(at' + bt'^2 + ct'^3 + l_j t'_j)] = 0; \tag{21}$$

where the summation extends over all plots, $t' = t - \bar{t}$, and t'_j implies the values of t' used at the j-th locality. Since $\sum t' = \sum t'^3 = 0$ over all plots, equation (21) simplifies as follows:

$$\sum yt' = 568b - (42l_1 + 30l_2 + 18l_3 + 6l_4 - 6l_5 - 18l_6 - 30l_7 - 42l_8). \tag{22}$$

TABLE 7

COEFFICIENTS OF LEAST SQUARES EQUATIONS FOR ESTIMATION OF OPTIMUM PLANTING DATE OF COTTON

Independent Variables

Plot Totals*	l_1	l_2	l_3	l_4	l_5	l_6	l_7	l_8	a	b	c
L_1 : 29.11	12	0	0	0	0	0	0	0	12	−42	155
L_2 : 82.85	0	12	0	0	0	0	0	0	12	−30	83
L_3 : 63.61	0	0	12	0	0	0	0	0	12	−18	35
L_4 : 27.07	0	0	0	12	0	0	0	0	12	− 6	11
L_5 : 50.55	0	0	0	0	12	0	0	0	12	6	11
L_6 : 46.00	0	0	0	0	0	12	0	0	12	18	35
L_7 : 62.65	0	0	0	0	0	0	12	0	12	30	83
L_8 : 45.76	0	0	0	0	0	0	0	12	12	42	155
G : 407.60*	12	12	12	12	12	12	12	12	96	0	568
B : −39.22*	−42	−30	−18	− 6	6	18	30	42	0	568	0
C : 2299.62*	155	83	35	11	11	35	83	155	568	0	6742

Solution for b and c

(1) $B + (7L_1 + 5L_2 + 3L_3 + L_4 - L_5 - 3L_6 - 5L_7 - 7L_8)/2 = -32.32 = 64b$

(2) $C - (155L_1 + 83L_2 + 35L_3 + 11L_4 + 11L_5 + 35L_6 + 83L_7 + 155L_8)/12 = -64.68 = 4096c/3$

(3) $G = 96a + 568c$

*G = Grand Total; $B = \Sigma\, y(t - \bar{t})$; $C = \Sigma\, y(t - \bar{t})^2$; Plot totals at Kawanda.

The coefficients for all of the equations and the requisite data from the Kawanda experiment are given in Table 7. If we multiply the L equations by the appropriate constants, as shown at the bottom of Table 7, and add to the b and c equations, we have at once the equations to estimate the values of b and c adjusted for locality effects,

$$b = -0.5050, \qquad c = -0.04737. \tag{23}$$

And then $a = \bar{y} - 568c/96 = 4.5261$.

From these final equations, we also see that b and c are independent and that the variances of b and c are

$$\sigma^2(b) = \sigma^2/64; \qquad \sigma^2(c) = 3\sigma^2/4096, \tag{24}$$

where σ^2 is estimated from the error variance, in this case 0.3606. Hence the estimated variances and standard errors of b and c are

$$\begin{aligned} s^2(b) &= 0.005634; \qquad & s(b) &= 0.07506 \\ s^2(c) &= 0.0002641; \qquad & s(c) &= 0.01625 \end{aligned} \tag{25}$$

These results indicate that both b and c are significantly different from 0.

A comparison of the adjusted mean yields based on the quadratic regression equation (20) and on the original equation (7) (Table 5) is given below.

ADJUSTED MEAN YIELD

	Date	(7)	(20)	Deviation
	1	5.691	5.839	−.148
	2	5.869	5.713	.156
	3	5.501	5.492	.009
	4	5.247	5.177	.070
(26)	5	4.372	4.767	−.395
	6	4.413	4.262	.151
	7	3.778	3.662	.116
	8	2.989	2.968	.021
	9	2.090	2.178	−.088
	10	1.391	1.294	.097
	Average*	4.247	4.247	.000

*The average is computed as $[d_1 + d_{10} + 2(d_2 + d_9) + 3(d_3 + \cdots + d_8)]/24$.

From (26), we see that there is no pronounced trend in the deviations, such as consistent positive or negative deviations at the early and late dates or in the middle. The only large deviation is at the fifth date, for which we previously noted the unexplained sharp drop. The chief difference between the two series of adjusted yields is that the maximum point is at the first date using the quadratic regression as compared to the second date for the original adjusted yields.

It might be advisable to check the adequacy of our quadratic prediction equation in estimating the adjusted yields. We find that the reduction in total sum of squares due to the quadratic regression is given by

$$(-32.32)(-0.5050) + (-64.6758)(-0.04737) = 19.3853. \tag{27}$$

Hence the remaining sum of squares for the other 7 degrees of freedom for dates is

$$21.7354 - 19.3853 = 2.3501 \tag{28}$$

This mean square is $2.3501/7 = 0.3357$, which is even less than the error mean square.

The planting date to give the maximum yield can be estimated by differentiating the estimating quadratic equation (20) with respect to t and equating the result to 0. This gives as a result the maximum planting date[1]

$$t_{max} = \bar{t} - b/2c = 0.17 \tag{29}$$

The variance of this estimate can only be approximated. If only the first order terms of the Taylor expansion of the differential of $b/2c$ is used, we find that

$$\begin{aligned} \sigma^2(b/2c) &= \frac{\sigma^2(b)}{4c^2} + \frac{b^2\sigma^2(c)}{4c^4} = \left(\frac{b}{2c}\right)^2\left[\frac{\sigma^2(b)}{b^2} + \frac{\sigma^2(c)}{c^2}\right] \\ &= \left(\frac{b}{2c}\right)^2\left[\frac{1}{64b^2} + \frac{3}{4096c^2}\right]\sigma^2, \end{aligned} \tag{30}$$

where σ^2 is estimated by s^2. Hence

$$s^2(b/2c) = 3.994; \qquad s(b/2c) = 2.00. \tag{31}$$

The standard error of $t_{\max}$ is the same as $s(b/2c)$, because $\bar{t}$ is fixed. Hence the maximum planting date ± two standard errors is given by 0.17 ± 4.00. As expected, the confidence interval is very large, indicating that the difference between the two maximum points in (26) is unimportant. The variance of the estimate of the optimum planting is large for two reasons: (i) There is a serious loss of information in making the adjustments for locality differences. For example, the coefficient of b is reduced from 568 to 64 in making the adjustments (Table 7). (ii) The optimum point comes before or near the first planting date used in the experiment. If the optimum planting date were near $\bar{t}$, so that b would be small, $s^2(b/2c)$ would be materially reduced.

(d) Effects of Lygus Infestation. Unfortunately this design does not furnish any means of determining if there is any linear decrease in the yield for the last planting at a given locality because of a Lygus migration from the earlier planted plots. If we include a constant in equation (20) to represent the difference between the yields of the first and third plantings at all localities, this constant proves to be the same as the linear regression coefficient, b, of yield on planting date, adjusted for localities. As indicated before, the quadratic effect of the Lygus migration can be measured by taking the differences between the yield for the second planting and the average of the yields for the first and third plantings. This quadratic effect was shown to be non-significant.

---

[1] This is a maximum only if c is negative.

TABLE 8

INDEX OF LYGUS DAMAGE PER COTTON PLANT FOR THE KAWANDA EXPERIMENT

Locality	1			2			3			4		
Date	1	2	3	2	3	4	3	4	5	4	5	6
	2.16	1.16	2.89	4.02	2.00	1.54	3.66	2.17	1.46	1.78	2.13	2.47
	2.49	1.69	2.22	3.20	3.35	1.65	3.64	1.91	1.56	3.59	3.04	2.27
	1.71	1.60	2.89	3.69	4.11	1.09	3.84	1.52	1.56	2.17	2.87	2.76
	1.62	1.33	1.67	3.00	4.55	2.50	3.73	2.11	1.53	2.26	2.31	3.20
Total	7.98	5.78	9.67	13.91	14.01	6.78	14.87	7.71	6.11	9.80	10.35	10.70
	23.43			34.70			28.69			30.85		

Locality	5			6			7			8		
Date	5	6	7	6	7	8	7	8	9	8	9	10
	3.47	4.24	2.34	3.16	3.25	2.04	2.62	2.37	0.78	1.71	0.89	1.16
	3.58	3.89	2.31	3.85	2.05	1.67	1.51	2.39	1.11	2.46	1.49	1.51
	3.24	4.26	2.00	3.78	4.05	1.73	2.20	2.04	0.82	1.89	2.40	1.16
	4.96	3.69	3.07	4.00	2.16	2.02	1.62	2.04	1.20	2.60	1.13	0.71
Total	15.25	16.08	9.72	14.79	11.51	7.46	7.95	8.84	3.91	8.66	5.91	4.54
	41.05			33.76			20.70			19.11		

In order to check if there was an adverse effect on the cotton yield by Lygus migration, an index of the Lygus damage per plant was determined for each plot of the Kawanda experiment. The Lygus damage data are given in Table 8. Using these data, the cotton yields (Y) were adjusted for the Lygus damage (X) in the following analysis of covariance, where $x = X - \bar{X}$ and $y = Y - \bar{Y}$.

TABLE 9

ANALYSIS OF COVARIANCE FOR THE KAWANDA EXPERIMENT

Source	Degrees of Freedom	Sy^2	Sxy	Sx^2
Interaction	7	1.0882	0.4357	3.0523
Dates (adj.)	9	21.7354	19.8782	36.7967
Localities (adj.)	7	177.6513	29.7493	27.0614
Replications	24	11.6219	−2.5228	5.4713
Error	48	17.3125	−2.3957	14.0422
Error + Dates (adj.)	57	39.0479	17.4825	50.8389
Error + Loc. (adj.)	55	194.9638	27.3536	41.1036

Error	47	16.9038	.3597
Error + Dates (adj.)	56	33.0360	
Dates (adj.)	9	16.1322	1.7925**
Error + Loc. (adj.)	54	176.7605	
Localities (adj.)	7	159.8567	22.8367**

**Significant at the 1% probability level.

From this analysis, we conclude that there was no significant over-all regression of yield on Lygus damage, as shown by the non-significant reduction in the error sum of squares by the use of X-variate (Lygus damage). Hence we conclude that there was no serious insect migration to the last-planted cotton at a given locality, when *only three planting dates* were used at each locality.

The regression coefficient, $b_{Y \cdot X}$, is found from the error row in the above table:

$$b_{Y \cdot X} = \frac{-2.3957}{14.0422} = -0.1706. \tag{32}$$

Each yield figure could be adjusted for the Lygus damage by use of the equation

$$Y(\text{adjusted}) = Y + 0.1706(X - \overline{X}). \tag{33}$$

Then the optimum planting date could be determined from these adjusted yields by use of the methods given above.

It would not be useful in general to bother with such a small adjustment, but we will illustrate the method which would be used. In Table 7, the locality yield totals (adjusted for Lygus damage) are

$$28.15,\ 83.82,\ 63.55,\ 27.38,\ 52.60,\ 46.81,\ 61.23,\ 44.07. \tag{34}$$

Similarly the adjusted values of B and C are -51.46 and 2261.44. Using these adjusted values, we obtain

$$b = -0.5965; \qquad c = -0.05113. \tag{35}$$

The variances of these estimates will be the same as before except that σ^2 is estimated as the average variance of $Y - b_{Y \cdot X}(X - \bar{X})$.[2] This average variance is $0.3597(1 + 1/96) = 0.3634$. There is no significant change in the value of either b or c.

ANALYSIS OF THE KABULA DATA

As stated previously, the same experimental design was used at the Kabula Experiment Station, except that the planting dates ranged from September 13 through January 17. The average yield for this experiment was 4.463 kgms./plot. No data were gathered on the Lygus damage. The analysis of variance for the Kabula data is given in Table 10.

TABLE 10

ANALYSIS OF VARIANCE FOR THE KABULA COTTON YIELD DATA

Source	Degrees of Freedom	Mean Square
Interaction	7	0.7999
Dates (adj.)	9	3.6573**
Localities (adj.)	7	26.5159**
Replications	24	1.6730
Error	48	1.1448

**Significant at the 1% significance level.

The unadjusted mean yields per plot for each planting date at Kabula, the mean yields per plot adjusted for localities by equation (7) and the estimated adjusted mean yields per plot based on the quadratic regression equation (20) are presented in Table 11.

---

[2] This is only an approximation to the best estimate of σ^2. The correct estimate would involve weighting each observation according to its use in the estimation of b and c.

TABLE 11

MEAN COTTON YIELDS FOR EACH PLANTING DATE AT KABULA

Date	t'	Unadjusted Mean Yield	Mean Yields Adjusted for Localities Equation (7)	Equation (20)	Equation (38)
1 (9-13)	−4.5	6.498	6.808	7.279	6.665
2 (9-27)	−3.5	6.412	6.997	6.479	7.151
3 (10-11)	−2.5	6.018	6.631	5.751	6.525
4 (10-25)	−1.5	4.816	5.661	5.096	5.428
5 (11-8)	−0.5	2.544	3.827	4.514	4.325
6 (11-22)	0.5	2.362	3.774	4.003	3.503
7 (12-6)	1.5	3.466	3.065	3.566	3.077
8 (12-20)	2.5	4.611	2.939	3.200	2.983
9 (1-3)	3.5	6.033	2.996	2.907	2.985
10 (1-17)	4.5	4.286	2.644	2.686	2.666
Average		4.463	4.463	4.463	4.463

Again we note that there is no significant decrease in the adjusted mean yield until we reach the fourth or fifth planting date. Some of the differences between the adjusted mean yields, using equation (7), and their standard errors are given below.

	Date	Mean Difference	Standard Error
	2-4	1.336	.610
(36)	4-5	1.834	.498
	5-10	1.183	.976

It would appear that the optimum planting date is somewhere between September 13 and October 11.

When the quadratic trend was fitted to the mean yields at the successive dates by equation (20), we obtained the following estimates of b and c ± their standard errors:

$$\begin{aligned} b &= -0.5103 \pm 0.1338 \\ c &= 0.0362 \pm 0.0290 \end{aligned} \tag{37}$$

We note that only the linear coefficient, b, is significantly different from 0. Also since b and c are of opposite signs, this regression curve will be concave upwards, indicating a minimum rather than a maximum at the point: $t = \bar{t} - b/2c$. Hence no optimum planting date can be estimated by use of the quadratic equation. Extrapolation would be impossible because the adjusted means estimated by the regression will increase on both tails when c is positive. The minimum point is not reached until $t = 12.55$; hence, Table 11 does not indicate the increasing yields at the later dates. However, we do see that the regression estimate is beginning to diverge quite noticeably at the first date.

It might be useful to investigate the possibility that equation (20) should be changed to include third and fourth degree terms, as follows:

$$y_{ti} = 4[a + bt' + ct'^2 + dt'^3 + et'^4] + p_{ti} \: . \tag{38}$$

The degree must be even if we are to have the downward trend at both ends. The estimates of b and d (adjusted for localities) are independent of c and e. However b and d are correlated with each other and so are c and e.

The estimating equations for b and d are:

$$\begin{aligned} 64b + 1072d &= -32.658 \\ 1072b + 31060d &= -299.868 \end{aligned} \tag{39}$$

The inverse of the coefficient matrix is

$$\begin{bmatrix} .03703545 & -.00127824 \\ -.00127824 & .0000763126 \end{bmatrix} \tag{40}$$

The estimated values of b and d ± their standard errors are:

$$\begin{aligned} b &= -0.82622 \pm 0.2059 \\ d &= 0.018862 \pm 0.00935. \end{aligned} \tag{41}$$

Hence we conclude that both b and d are significantly different from 0.

Similarly for c and e, the estimating equations are:

$$\begin{aligned} (4{,}096c + 84{,}736e)/3 &= 49.441 \\ (84{,}736c + 2{,}038{,}528e)/3 &= 326.086. \end{aligned} \tag{42}$$

The inverse of the (c, e) coefficient matrix is:

$$\begin{bmatrix} .00522869 & -.000217342 \\ -.000217342 & .0000105060 \end{bmatrix} \tag{43}$$

Hence the estimated values of c and e with their standard errors are:

$$\begin{aligned} c &= .18764 \pm .07739 \\ e &= -.0073197 \pm .003467 \end{aligned} \tag{44}$$

The adjusted mean yields estimated by the quartic regression curve (38) are also given in Table 11. The agreement between these estimated mean yields and those estimated from equation (7) is quite good. The estimated mean yields (by quartic regression) are beginning to drop off at both ends, and there are no pronounced series of plus or minus deviations as with the estimated mean yields found by using only the quadratic regression. The remaining date sum of squares after fitting equation (38) is 4.6986 with 5 degrees of freedom, giving a mean square of 0.9397, which is smaller than the error mean square.

The problem of estimating the optimum planting date is complicated by the fact that there are two maximum points and one minimum point. These points are solutions of the equation

$$b + 2ct' + 3dt'^2 + 4et'^3 = 0. \tag{45}$$

The maximum point with the largest average adjusted yield is at $t' = -3.665$, which is slightly before the second date (September 27). No attempt has been made to estimate the standard error of this estimated optimum planting date.

RELATIVE EFFICIENCY OF THIS EXPERIMENTAL DESIGN

If there were no adverse effects from the Lygus migration to the newly planted cotton, it would be advisable to use as many planting dates as possible at each locality. The analysis of the Kawanda data seems to indicate that more than 3 successive planting dates might have been used at a given locality without incurring any serious insect interference. We shall consider the relative efficiency of the following field designs:

(i) The present "staggered" 3-date plan.

(ii) A staggered 4-date plan, with only 6 locations as follows:

Location	Dates	Location	Dates
1	1,2,3,4	4	5,6,7,8
2	2,3,4,5	5	6,7,8,9
3	3,4,5,6	6	7,8,9,10

(Notice that the sequence (4, 5, 6, 7) is omitted in order to have a design which would have the same number of replications per date as for the 3-date plan.)

(iii) A balanced incomplete blocks design with 3 planting dates per block, which requires 30 blocks with each planting date being replicated 9 times. (The same number of plots could not be planted at each location if this design were used.)

(iv) A complete blocks design with all 10 planting dates at each locality, the locality being a complete block.

Under designs (i) and (ii), 96 plots would be used, but only 90 plots would be used for (iii) and only 80 plots for (iv). Hence the expected error variances will have to be adjusted to an equal number of plots. The most efficient of these designs would be the complete blocks design (iv), because no adjustments for localities would be required in determining the average yield for a given planting date. The balanced incomplete design (iii) has been shown to be 74% as efficient as the complete blocks design (iv); that is, 4 replications of each planting date using (iii) would be required to give the same accuracy as 3 replications using (iv).

The relative efficiency of design (i) will also be assessed with respect to (iv) by comparing the average variances of the difference between the mean yields at successive planting dates. If the error variance per plot is designated as σ^2, then the variance of the difference between *any* two date means for the complete blocks design would be $2\sigma^2/8 = .25\ \sigma^2$. This variance can be compared with any of those in Table 6 by adjusting for the different number of total plots used in the two designs. Since 96 plots were used in the "staggered" 3-date plan (i), the variances in Table 6 should be multiplied by $96/80 = 1.2$ to put them on the same basis as that of the complete blocks design.

The lowest variance in Table 6 is that for the difference in yield between dates 5 and 6, $0.21652\ \sigma^2$. For comparison purposes, we multiply this variance by 1.2, giving $.26\ \sigma^2$. Hence this comparison is 96% efficient as compared with the same comparison if design (iv) were used. If we omit the comparisons for the first two and the last two planting dates, the average of the efficiencies for planting dates separated by one time interval (a fortnight) would be 94%. The comparisons for the first two and the last two dates are quite inefficient, because these planting dates are not used at many locations under design (i); the relative efficiency is only 48%. The average efficiency for all nine of the one fortnight comparisons is 84%. The average efficiencies for all the comparisons of average yields given in Table 6 are presented in Table 12 below. It should be noted that these efficiencies are based on the premise that σ^2 is the same for all designs. If the increase of block size to 10 plots per block in design (iv) results in an increase in σ^2, the relative efficiencies would be higher than those given in Table 12. The incomplete blocks design (iii) would have the same block size; hence,

σ^2 for this design would be expected to be the same as for the 3-date plan (i). The relative efficiency of (i) as compared to (iii) would be $1/.74 = 1.35$ times as great as the efficiencies given in Table 12.

As stated above, we also consider the efficiency of the 4-date plan (ii), assuming that σ^2 would be the same for this design as for (i), even though only 6 instead of 8 localities are considered. The least square equation for (ii) was set up and a variance table such as Table 6 was then constructed. The relative efficiencies of this design (ii) compared with the complete blocks design (iv) are also presented in Table 12. It might be mentioned that a balanced incomplete blocks design with 4 planting dates per block can be constructed, using 15 blocks. This balanced incomplete blocks design is 83% as efficient as the complete blocks design, assuming σ^2 does not increase with the increase of block size for the complete blocks design.

TABLE 12

AVERAGE PERCENTAGE EFFICIENCIES OF THE 3 AND 4-DATE PLANS (i and ii) COMPARED WITH A COMPLETE BLOCKS DESIGN (iv)

	Number of Fortnights Separating Planting Dates									
Design	1	2	3	4	5	6	7	8	9	Avg.
i*	94	65	46	36	30	25	22			57
ii†	106	89	68	52	44	41	34			74
i†	84	61	43	34	28	23	20	17	14	47
ii†	93	80	63	48	41	35	31	26	21	60

*Omitting the first and last planting dates.
†Using all planting dates.

From Table 12, we see that the 3-date plan (design i) is only about 3/4 as efficient as the 4-date plan (design ii). That is we could have obtained about the same variances of the adjusted mean differences with 4 replications at each of the 6 locations for (ii) as with the 4 replications at each of the 8 locations for (i).

It might be noted that in this particular experiment, the most important comparison for the determination of the range of the optimum planting dates was that between dates 1 and 5. The efficiency of the estimate of the difference in the mean yield between these two dates is only 30% for the 3-date plan (i). Actually since the optimum planting date could have come before the first planting date used in the experiments, this range should have been about twice as large; the efficiency of this estimate would be even lower than 30%. Hence we can conclude that the 3-date design is a very inefficient design for determining the optimum planting date.

The efficiency of the determination of the optimum planting date can also be estimated from the variances of b and c in the quadratic regression equation (20). The variances of b and c for the 3-date plan (using 96 plots) are given in equation (24).

The variances of b and c for the four designs (i, ii, iii and iv), based on 80 plots, are the following coefficients of σ^2:

		(i)	*(ii)*	*(iii)*	*(iv)*
(46)	$\sigma^2(b)$	.01875	.01000	.00214	.00152
	$\sigma^2(c)$	.000882	.000514	.000320	.000237

These results show that the 3-date plan (i) is only about 8% efficient in the estimation of b and 27% efficient in the estimation of c as compared to the complete blocks design, if we can assume that σ^2 will not increase with the latter. When compared with the incomplete blocks design, the efficiency is about 1.5 times as great. Comparing the 3 and 4-date plans, we note that the efficiency is almost doubled for the 4-date plan.

From equation (30), we see that the variance of the optimum planting date (if the regression is quadratic) is

$$\sigma^2(t_{max}) = \frac{1}{4c^2}\left[\sigma^2(b) + \frac{b^2}{c^2}\,\sigma^2(c)\right], \tag{47}$$

where $\sigma^2(b)$ and $\sigma^2(c)$ are found by multiplying the constants given in (46) by σ^2. In order to increase the efficiency of the experiment, we must either reduce σ^2, redesign the experiment so as to obtain lower coefficients for $\sigma^2(b)$ and $\sigma^2(c)$, or plan the experiment so as to have the optimum planting date fall near the middle of the dates used in the experiment (make b small).

If we do set up the experiment so that b is expected to be small, then the efficiency of the experiment largely depends on $\sigma^2(b)$. From (46), we see that $\sigma^2(b)$ is very large for the "staggered" designs as compared to the randomized blocks designs (iii and iv). It has been suggested that we might increase the efficiency of the "staggered" 3-plan design by redesigning the experiment. For example we might use only the following sequences of dates—1, 2, 3; 2, 3, 4; 7, 8, 9; 8, 9, 10—planting each sequence at two locations. Although $\sigma^2(c)$ is reduced from $3\sigma^2/4096$ to $3\sigma^2/7168$, $\sigma^2(b)$ remains at $\sigma^2/64$ (for 96 total plots). This design is not recommended for two reasons: (i) It does not improve the efficiency of the estimation of b, which is the chief contribution to the inefficiency of the "staggered" designs, (ii) There should be some estimate of the yield for all planting dates, especially since the hypothesis of a quadratic regression of yield on planting date may be false. If some other regression equation is used, the reduction in the variance of c may be offset by an increase in the variance of some other estimate.

If we were to use the sequences—1, 2, 3; 3, 4, 5; 5, 6, 7; 7, 8, 9—all planting dates, except the tenth, would be represented. However $\sigma^2(b)$ would still be $\sigma^2/64$ and $\sigma^2(c)$ would be increased to $3\sigma^2/3904$.

SUMMARY AND CONCLUSION

This paper presents a new type of experimental design, the "staggered" design, for use with experimental material which can have but few consecutive plantings at a given locality. Two experiments involving the determination of the optimum planting date of cotton have been conducted in British East Africa. The "staggered" design was used here because of a fear that an insect, *Lygus simonyi*, would tend to transfer from earlier planted plots to newly planted ones, hence distorting a proper assessment of the relationship between cotton yield and planting date. At one of the experiments, an index of the Lygus damage was determined for each plot. If this index truly reflected the Lygus damage, it appeared that there was no important migration for planting dates separated by four weeks or less.

We have shown that the 3-date "staggered" design is decidedly inefficient in estimating the optimum planting date. The adjustments for locality effects are so great that there tends to be a very long range of indeterminacy of the optimum planting date. The following suggestions are offered to improve the efficiency of the estimation of the optimum planting date under conditions of uncertainty with respect to insect migration:

(i) Test the adequacy of the Lygus damage index as a true indicator of the infestation by the Lygus bug. If this index is reliable, or can be made reliable, then it appears that more planting dates should be used at each locality and that adjustments should be made for the Lygus damage by covariance techniques rather than by use of the "staggered" design.

(ii) Even if the index is shown to be inadequate, especially when more than 3 planting dates are used at each locality, we would advise at least trying 4 successive planting dates at each locality.

(iii) A great improvement would result if the localities did not differ so widely in their fertility. This "staggered" design would be much more efficient if the locality differences were not so pronounced as in these African experiments.

(iv) It is advisable to plan the experiment so that the optimum planting date is near the middle date used in the experiment.

(v) Ordinarily the experimenter hopes to secure some information on the correct allocation of experimental material as to changes in the number of locations and replications within locations. In this case the small date-location interaction as compared to the date-replication interaction leads us to infer that one could not lose any information by using fewer localities and more replications at each locality. Since this result was obtained in both experiments, we feel that the use of more planting dates at a given locality with fewer localities being used probably would not materially alter the error variance.

ON THE USE OF PARALLEL OR NONPARALLEL SYSTEMS OF TRANSFORMED CURVES IN BIOASSAY: ILLUSTRATION IN THE QUANTITATIVE COMPLEMENT-FIXATION TEST*

William R. Thompson

Division of Laboratories and Research,
New York State Department of Health, Albany

A method of moving-average interpolation has been presented elsewhere [1,2a] to provide a sound basis for quantal bioassay independent of any assumption as to the precise form of the dosage-response curve. As the method is relatively simple in application as well as in assumptions involved, it may be regarded as a *basic method* in the sense [1] that it may be preferred except in any given situation where the use of some other method can be justified by improved precision or by permitted technical economies. Moreover, it was shown that the Kärber [3–6] and Reed-Muench [7] methods (designed to serve a like purpose) have serious defects, and that modifications introduced to eliminate these defects led directly to the moving-average interpolation method. The characteristics of various prevalent curve-fitting methods were also discussed with emphasis upon certain features usually ignored, especially with regard to systems of curve fitting by minimizing the sum of weighted squared deviations and use of *adjusted* or *fictitious* points. However, the writer was careful to emphasize that he has no intention of discouraging use of the curve-fitting methods, which appear to have found many valuable applications.

Now with regard to the systems of curves that may be used to approximate either the dosage-response curve or what has been designated previously [1] the *fundamental* (log dose, response) *curve*; it is often considered essential to a valid system of bioassay that the log dose-response curves employed should all have the same shape. This is equivalent, of course, to providing that the curves be transformable to a system of parallel straight lines with log dose as abscissa and some univariant function of response (for example, the probit or the logit) as ordinate. The present purpose is not only to examine conditions under which such a stipulation may be justified, but in striking contrast to suggest other possible conditions where such a stipulation would be *invalid* yet where valid systems for bioassay by curve fitting may exist, and to provide illustrations of each of the two situations drawn from investigations of the quantitative complement-fixation reactions. Incident to this development certain relations, familiar in physical chemistry, are used to portray a part of the nature of these biochemical reactions and to suggest possible features of reactions *in vivo*.

---

*Presented at the Annual Meeting of the American Statistical Association, Biometrics Section, New York, N. Y., December 29, 1947.

Accordingly, suppose that we are dealing with a *quantal* (all or none response) assay system. In a given case let D be the dose of an agent C allowed to act (possibly in some cases modified by introduction also of an amount V of an inhibitor F) on a given number n of *subjects* (individuals treated), and let r be the number of these that *respond critically*, and $p = r/n$. Then, under given conditions with V fixed (possibly zero), assume that we have a continuous univariant function of x,

$$y = f(x) = f^*(x, a, b, \cdots) \tag{1}$$

where $0 \leqq y \leqq 1$ for $x > 0$, which may be used in conjunction with some prescribed fitting process applied to certain parameters (a, b, etc.) so that the (x, y) curve will satisfactorily approximate the set of *observed points* (D, p) obtained in a bioassay. Let K be the estimate of median-effective dose defined by $f(K) = 0.5$. In order to simplify the discussion without loss of essential generality, assume that the critical response is so defined that $f(x)$ is an *increasing* continuous function of x. Also, let L denote $\log x$, and let $\phi(L)$ be identical with $f(\text{antilog } L) \equiv f(x)$. Then $y = \phi(L)$ describes the approximate *fundamental curve* with y as ordinate and L as abscissa.

Now, let $X(x)$ and $Y(y)$ be increasing continuous functions respectively of x and y. Further assume that these functions may be defined so that

$$X(x) = a + b.Y(y) \tag{2}$$

where a and b are parametric constants (obviously $b > 0$). Further restrict the Y-function by the identity,

$$Y(0.5) \equiv 0, \tag{3}$$

which may be done without loss in essential generality. Accordingly, by the definition of $K = x$ if $y = 0.5$, we have

$$X(K) = a. \tag{4}$$

In practice we may wish to relax these restrictions on $Y(y)$ which, however, facilitate the present discussion. Thus replacing $Y(y)$ in (2) by $Y_0(y) = \pm(Y(y) + \Delta)$, and replacing a and b respectively by a_0 and b_0 and applying the same curve-fitting process should result in the evaluations of these parameters as $b_0 = \pm b$ and $a_0 = a - b_0 \cdot \Delta$ with the resultant value for K identically the same, $X(K) = a = a_0 + b_0 \cdot \Delta$. An example would be use of $Y_0(y) =$ probit y which is equivalent to use of $Y(y) =$ the normal equivalent deviation of y (for unit standard deviation), which may be denoted by N.E.D. of y; in this case $b = b_0$ and $\Delta = 5$. Another example would be $Y_0(y) = \log [(1 - y)/y]$ instead of $Y(y) = \log [y/(1 - y)]$; then $b_0 = -b$ and $a_0 = a$ as $\Delta = 0$.

In a personal communication Berkson has informed the writer that he plans to change his definition of logit y to equal $\log [y/(1 - y)]$ rather than minus that value.

Obviously, in any case relation (2) has the advantage of permitting a fit of a straight line to experimental points $[X(D), Y(p)]$ graphically by inspection or by some least square technic with or without a weighting system or by some other specified curve-fitting process such as those mentioned above. Of course, relation (2) is equivalent to

$$(5) \quad Y(y) = a' + b'.X(x) \text{ for } b' = 1/b \text{ and } a' = -ab'.$$

Use of the form (2) or (5) is not intended here to indicate any prejudice as to the direction in which any deviations are to be measured nor as to the curve-fitting technic to be used. It seems preferable to suspend judgment on such issues temporarily, since the principal object of the present discussion is to indicate what kind of systems of curves may be considered appropriate to given situations, especially with regard to whether or not in (2) we are to use a family of parallel straight lines to fit two or more sets of assay data; i.e., to make a simultaneous fit with the same value of b but individual values of a for each set, or possibly with a prescribed value of b. Moreover, as shown previously (1), the weighting systems commonly used for cases where n is relatively small (say 20 or less) take into account only the supposed principal source of error, biologic variation in individual resistance to the agent in question and sampling error corresponding to use of n individuals in a given case under test; the systems may be inappropriate for cases where n is large, for example, in the quantitative complement-fixation tests where n is about 100 million or more.

The pioneer work in application of transformations as in (2) to data as obtained in bioassay appears to have been that of von Krogh [8], who employed

$$(6) \qquad X(x) = \log x \quad \text{and}$$

$$(7) \qquad Y(y) = \log \frac{y}{1 - y}$$

to convert observed points to approximately linear form, where y is the degree of hemolysis induced in sensitized red blood cells by a given amount x of complement under given conditions. This was used by Wadsworth, Maltaner, and Maltaner [9] and others [10–18, 2] in quantitative studies of complement fixation and in complement titration, which furnished the basis for the quantitative tests for specific antibody content in blood or other body fluids. Thus, the logistic function converted to logarithmic form

$$\log x = \log K + h.\log \frac{y}{1-y} \tag{8}$$

was employed[1], where K is the estimated amount of complement required under otherwise the same conditions to induce 50-per-cent hemolysis (the median-effective dose) and h is a parametric constant. Under these conditions in (2) we should have $b = h$ and $a = \log K$.

The logistic curve has been applied thus to other forms of bioassay by Wilson and Worcester [20,21] and by Berkson [22,23], with different curve-fitting methods. The former have also considered [24] general relations permitting use of other curve forms. They stress the importance of recognizing that no one curve form should be assumed suitable for use in all situations, which may involve different laws of biologic reaction of various animals to various biologicals. Some of the possible types of difficulty that may be encountered have been discussed elsewhere [1,2a].

The principal rival of the logistic as approximation to the *fundamental* (log dosage-response) *curve* is the *integrated normal curve* (taking probit y for $Y_0(y) = 5 +$ the N.E.D.). Winsor [25] has shown that either may be fitted to the other so well over the ranges usually employed in bioassay that it would ordinarily be difficult to discriminate between them on the basis of goodness of fit to experimental data and usefulness in estimation of median-effective dose. Accordingly, it would seem naive to use an experimentally observed correlation between two methods, one using the logistic and the other the integrated normal curve, as an argument in favor of either method in a given situation. Likewise it should be expected that the integrated normal curve could be used instead of the logistic to represent the hemolysis data, and this has been done by Ipsen [26]. Bliss and Cattell [27a] in an extensive review cite Ipsen's work, but apparently 1914 is incorrectly given as its publication date in place of 1941.

Now consider certain special conditions (denoted by roman numerals) that may or may not be assumed to apply to given reaction systems.

---

[1] See footnote on page 208.

I. Two or more different preparations of essentially the same reagent except for possible differences in concentration are to be compared in tests where x, the amount used in a given test, is variable; but all other conditions are planned to be essentially the same (for example, an amount of diluent is used to make the total volume of material introduced essentially the same).

Under condition I, if K_i is the median-effective dose of the i-th preparation and x_i is the dose of this preparation employed ($i = 1, \cdots, N$), and we take $x_i/K_i = x_j/K_j (j = 1, \cdots, N)$, then the expected effects (y_i, y_j) are the same. Accordingly, by relations (2) and (4), we have

$$(9) \qquad \begin{aligned} X(x_i) &= X(K_i) + b_i . Y(y_i) \qquad \text{and} \\ X(x_j) &= X(K_j) + b_j . Y(y_j), \end{aligned}$$

where $Y(y_i) = Y(y_j)$. Obviously, if $X(x) = \log x$, then $b_i \equiv b_j$; and it is readily demonstrated that if $b_i = b_j$ identically (for every possible i and j under the circumstances and every possible x/K ratio) then $X(x)$ must be a linear function of $\log x$; and $\log x$ obviously will serve as well instead for our purposes.

Now, this situation, where the use of $X(x) = \log x$ in relation (2) leads to a system of parallel straight lines (with suitable choice of the Y-function), appears to be at least approximately realized in many types of bioassays, even under conditions different from those of I. However, this has led occasionally to incautious statements which seem to proscribe any possibility of another satisfactory basis for bioassay. Accordingly, consider some alternative conditions that may throw light on this issue.

II. Assume we are dealing with a single preparation of the active agent C, the amount x_i being used in the i-th test ($i = 0, 1, \cdots N$), but also introduced is an amount V_i of an inhibiting agent F (except that $V_0 = 0$). It is assumed that this amount of inhibitor reacts quantitatively with a proportional amount of C to take it out of action (essentially as if firmly bound in chemical combination), but that x_i is so taken as to leave some excess of active agent C, the excess being $x_i - g \cdot V_i$ and g is a constant. Further assume that the combined material (compounded of C and F) does not influence the reaction appreciably. Obviously, then $K_i = K_0 + g \cdot V_i$; and if $x_i = x_0 + g \cdot V_i$ then $y_i = y_0$ in any case. Accordingly, it is readily demonstrated that it is *necessary and sufficient* in (2) under condition II that $X(x)$ be a linear function of x if a system of parallel lines is attained. The argument proceeds as under (9) in even simpler manner but with x instead of $\log x$ as the principal form of $X(x)$. The latter is clearly impossible under condition II, $X(x) = \log x$ *will not serve* in (2) to give a family of parallel straight lines.

This may seem a rather extreme, perhaps an unnatural case. However, if the agent C resembled hydrogen-ion from hydrochloric acid, inhibitor F resembled hydroxyl ion from sodium hydroxide, and the compound formed resembled water we might find the conditions of II approximately realized. However, we are led at once to consider the corresponding situation where C and F do not react completely but some of each remains dissociated. This is the alternative condition following.

III. Assume that we are dealing as in II with a given preparation C introduced in the amount x_i with an amount V_i of inhibitor F also introduced, $V_0 = 0$ and $i = 0, 1, \cdots, N$. Assume that C and F may react in accord with the stoichiometric equation

$$F + C = FC \tag{10}$$

which we assume to be reversible. Assume that this reaction tends closely to approximate equilibrium conditions where $[C]_i$, $[F]_i$, and $[FC]_i$ may be used to denote the amounts respectively of C, F, and FC present at equilibrium, and that these amounts are expressed in equivalent units in accord with relation (10) and are related in accord with

$$\frac{[F]_i . [C]_i}{[FC]_i} = k \tag{11}$$

where k is a constant times the total volume in which the agents are able to act, which, in turn, is assumed constant and hence k is constant here. This is the same relation found by Northrop and Hussey to apply to trypsin and antitrypsin of the blood [28,29]. Assume that x and v are measured in the equivalent units (where $v = g \cdot V$ and g is a constant as under condition II), then

$$x_i = [C]_i + [FC]_i , \qquad \text{and} \tag{12}$$

$$v_i = g . V_i = [F]_i + [FC]_i . \tag{13}$$

Of course, as $v_0 = 0$, $x_0 = [C]_0$. Further assume that in any case under III the resultant y_i depends only upon $[C]_i$, the amount of *active* agent C.

Now, examine the characteristics of III by focusing attention on the special case,

$$[C]_i = x_0 \qquad \text{for } i = 0, 1, \cdots, N. \tag{14}$$

Accordingly, $y_i = y_0$ for every i, whence (11) gives

$$[F]_i = [FC]_i . k/x_0 , \qquad \text{and} \tag{15}$$

$$\begin{aligned} v_i &= [FC]_i .(1 + k/x_0) \\ &= (x_i - x_0)(1 + k/x_0). \end{aligned} \tag{16}$$

Accordingly, $v_i - v_j = (x_i - x_j)(1 + k/x_0)$ provided that $y_i = y_j = y_0$. Thus any endpoint y_0 may be used to furnish a measure of differences in amount v of total inhibitor present, since this is proportional to the corresponding differences in the amount x of reagent required to produce the specified response y_0, but *the factor of proportionality* $(1 + k/x_0)$ *will depend upon the endpoint chosen.* For many purposes it seems desirable to choose $y_0 = 0.5$ as the endpoint (this usually being about the point at which the corresponding x may be determined with greatest precision); and then, obviously, it is convenient to express x in terms of an arbitrary unit (13) taken so that $K_0 = 1$. Then

$$v_i = (K_i - 1)(1 + k). \tag{17}$$

Of course, k is an unknown constant. However, if V is given in terms of volume of inhibitor preparation introduced, then $g = v/V$ is the concentration of inhibitor in the preparation; and, in default of knowledge of k, we may use the *increment ratio titer*

$$T = (K_i - K_j)/(V_i - V_j) \tag{18}$$

which equals g divided by the unknown factor $(1 + k)$. This reduces to $T = (K_i - 1)/V_i$ for $V_j = V_0 = 0$; but it may be preferable to have both values of V greater than zero to avoid discrepant results that may be more likely in the region where little or no inhibitor is present.

Now, consider the possible influences of these conditions on the form of the curve, $f(x)$, that may be useful in the bioassay by relations (1) and (2) with suitable definition of $X(x)$ and $Y(y)$ and restriction of the range of y to that proposed for use in such assays. Relation (16) gives for the case, $y_i = y_0 = y$ (taken arbitrarily in the given range),

$$x_i = x_0 + v_i/(1 + k/x_0), \qquad \text{whence} \tag{19}$$

$$x_i/x_0 = 1 + v_i/(x_0 + k). \tag{20}$$

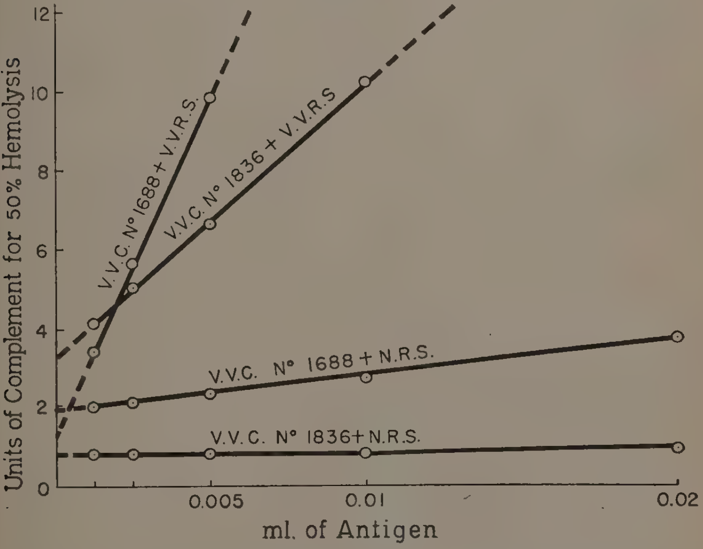
IIIA. Accordingly, if k is negligibly small relative to x_0, then (19) gives approximately $x_i = x_0 + v_i = x_0 + g \cdot V_i$, the same as condition II; and no transformation with $X(x) = \log x$ will yield a set of parallel straight lines in (2) respectively corresponding to various values of v_i, but $X(x) = x$ instead will yield that result approximately.

IIIB. On the other hand, if v_i is negligibly small relative to $x_0 + k$ or if x_0 is negligibly small relative to k, then x_i/x_0 is approximately constant for any bioassay (set of tests in which $v_i = v$, a constant, but x_i is varied to give a consequent response y correspondingly variable). Accordingly, $X(x) = \log x$ would result approximately in a set of parallels in (2) by suitable choice of $Y(y)$.

IIIC. However, there is a vast possible middle ground wherein v_i is not negligible relative to $x_0 + k$, nor either of the latter negligible relative to the other. On this ground, neither x nor $\log x$ may be used for $X(x)$ to provide a system of parallels in (2), yet by (16) or the derived relations (17) or (18) we see that *there exists an equally valid basis for assay* in all cases; results based on amount of agent required for a specified effect (y_0) furnish a sounder foundation for assay than any based on suppositions of a system of parallel straight lines to be found by suitable transformations which may not exist, or may be elusive. In an *Appendix* it is shown that they do not exist.

As in the case of any general statement, so with the denial of existence of any other valid basis for bioassay than that valid under condition I, a single example cited to contradict the general statement is all that is needed to disprove it. More than one contradicting hypothetical example has been exhibited. It is not claimed that these represent anything more than a first approximation to any real situation; but condition III comes fairly close to describing conditions that are found in several of the quantitative complement-fixation tests with C representing complement and F representing an antigen-antibody complex. In the case of the reactions with tuberculous immune serum and antigen [10], if A denotes the antigen and B the antibody, then the reactions observed closely approximate what would be expected if $A + B = AB$ and $AB + C = ABC$ were the stoichiometric relations involved, with AB in place of F as the antigen-antibody complex which inhibits the action of complement. Moreover, the system behaves approximately as if not only the second reaction but also the first followed a *mass law* resembling that given in relation [11], as may be noted in figure 3 in the cited article [10]. Linear relations are found in nearly all if not all these systems between the amount K of complement required for 50-per-cent hemolysis and either the amount of antigen or of antibody in the presence of the optimum amount of the other. An example is given in Figure 1 taken from an article by Doctor Rice on the vaccinia virus system [17]. In Table 2 of the same article are given values of h (there called $1/n$) for relation (8), which is the slope b as given in relation (2), as found experimentally under a variety of circumstances with different antigen-antibody systems. In the presence of both serum and antigen, the slope h may be seen to tend to increase as the amount of complement used to bring about a conveniently measureable reaction increases *only* in the case of the vaccinia virus system; in all others it decreases or remains about the same. The last situation is exhibited strikingly with the pneumococcus system. Approximately constant h is found for the higher preliminary incubation temperature but not for the lower, as indicated in Doctor Rice's table. Constant h was also found with the egg albumen system studied by Maltaner and Maltaner [12].

FIGURE 1



The ordinate is K, the median-effective dose of complement, shown as an approximately linear function of amount of vaccinia virus antigen (V.V.C.) present with the optimal amount of pooled antiserum (V.V.R.S.) or of normal serum (N.R.S.) of rabbits.

This figure is reproduced, by permission, from the *Journal of Immunology*, 1946, *53*, 225-236 (Williams and Wilkins Co.).

In an appendix it is shown that under condition III there exists no transformations, $X(x)$ and $Y(y)$ in (2), that yield a system of parallels. On the other hand, it is readily seen that we might measure variation in the amount of one inhibitor even in the presence of other inhibitors under certain similar conditions provided the latter inhibitors were present in approximately constant amount. This may be actually the case in some living systems and suggests a possible explanation of the approximation to condition IIIB often observed.

APPENDIX

THE POSSIBILITY OF SOLUTIONS UNDER CONDITION III

Assume tentatively under condition III that there exists a continuous function of x, $X(x)$, for $x > 0$; that its first derivative $X'(x)$ exists and is continuous, positive, and finite for $x > 0$; and that the X-response curves for different assays under the given conditions all have the same shape. The last is equivalent to assumption that a function $Y(y)$ exists as previously defined such that (2) holds with b constant for all assays and a the sole parameter distinguishing the corresponding parallel straight transform lines. Then $b > 0$ and the first derivative $Y'(y)$ exists and is continuous, positive, and finite for $0 < y < 1$.

To simplify the notation let $v = v_i$ and x_v denote the corresponding x_i under III which produces the response y in the presence of an amount of total inhibitor equal to v. Then, as before, x_0 is the value of x_v for $v = 0$, and (19) gives

$$\begin{aligned} x_v &= x_0 + v/(1 + k/x_0), \qquad \text{and} \\ K_v &= K_0 + v/(1 + k/K_0) \end{aligned} \tag{A1}$$

since $x_v = K_v$ if $y = 0.5$. Accordingly, (2) gives

$$X(x_v) = a_v + b \cdot Y(y) = X(K_v) + b \cdot Y(y) \tag{A2}$$

since $Y(0.5) \equiv 0$ and thus $a_v = X(K_v)$ is a parametric constant and b is a constant. Accordingly,

$$\frac{dX(x_v)}{dY(y)} = \frac{dX(x_v)}{dx_v} \cdot \frac{dx_v}{dY(y)} = X'(x_v) \cdot \frac{dx_v}{dY(y)} = b, \tag{A3}$$

a constant; and therefore

$$X'(x_v) \cdot \frac{dx_v}{dx_0} = X'(x_0). \tag{A4}$$

Now, (A1) gives $dx_v/dx_0 = 1 + kv/(x_0 + k)^2$.

Let $\Delta x = x_v - x_0$; then (A1) gives

$$\Delta x = v/(1 + k/x_0) \text{ whence } v = (1 + k/x_0)\cdot\Delta x,$$

whence (A4) gives

$$\text{(A5)}\qquad X'(x_0 + \Delta x)\cdot[1 + k\cdot\Delta x/x_0(x_0 + k)] = X'(x_0)$$

$$\text{(A6)}\qquad \therefore\ \frac{X'(x_0 + \Delta x) - X'(x_0)}{\Delta x} = \frac{-k\cdot X'(x_0 + \Delta x)}{x_0(x_0 + k)}.$$

Accordingly, we obtain the linear differential equation,

$$\text{(A7)}\qquad (x^2 + kx)\cdot X''(x) + k\cdot X'(x) = 0,$$

which may be considered as linear of first order in $X'(x)$. By well-known methods [30,31] this yields the general solution $X'(x) = C_1(1 + k/x)$ and hence

$$\text{(A8)}\qquad X(x) = C_1(x + k\cdot\log x) + C_2$$

where C_1 and C_2 are arbitrary constants except that $C_1 > 0$ in the present case as $X'(x) > 0$ for $x > 0$.

Accordingly we have in (A8) a necessary condition if $X(x)$ exist under relations (A1) and (A2). Obviously, we may take $C_2 = 0$ without loss in essential generality, and likewise take $C_1 = 1$. Thus, if there is any solution to (A1) and (A2) for $X(x)$, then we may use

$$\text{(A9)}\qquad X(x) = x + k\cdot\log x$$

as well as any for our purpose. Now, (A2) gives

$$\text{(A10)}\qquad X(x_v) - X(K_v) - X(x_0) + X(K_0) = 0.$$

Define ω equal to the left member of (A10) with $x + k\cdot\log x$ substituted for $X(x)$. Then

$$\text{(A11)}\qquad \omega = x_v - K_v + k\cdot\log\frac{x_v}{K_v} - x_0 + K_0 - k\cdot\log\frac{x_0}{K_0}$$

whence (A1) gives

$$\text{(A12)}\qquad \omega = \frac{k\cdot v(x_0 - K_0)}{(x_0 + k)(K_0 + k)} - k\cdot\log\frac{1 + v/(K_0 + k)}{1 + v/(x_0 + k)}.$$

Obviously, for ω to vanish identically, we have a solution if $k = 0$ (approximated in condition IIIA) and then we take $X(x) = x$.

However, *in any case* for $\omega \equiv 0$, then differentiating both sides of (A12) with respect to v gives

$$\text{(A13)} \qquad 0 = \frac{k(x_0 - K_0)}{(x_0 + k)(K_0 + k)} - \frac{k(x_0 - K_0)}{(x_0 + k + v)(K_0 + k + v)}$$

whence: either $k = 0$, or $x_0 = K_0$ or $v = 0$.
Thus, if $k \neq 0$ there is *no solution* to (A1) and (A2). It is readily verified that

$$\text{(A14)} \qquad \left| \frac{\omega}{K_0} \right| < \frac{k \cdot v^2}{2K_0} \cdot \left| \frac{1}{(K_0 + k)^2} - \frac{1}{(x_0 + k)^2} \right|$$

which indicates the conditions for which an approximate solution to (A1) and (A2) exists.

The lack of such solutions under condition III means only that transformations of assay curves to straight parallel lines by means of some functions $X(x)$ and $Y(y)$ do not exist. Valid conditions for assay relative to some fixed reference point (y_0) have been shown to exist *always* under condition III.

---

[1] In a general discussion of bioassay in another article [1] a different notation was employed as there more convenient, but a key was given (page 116, footnote 2) to other notations including that used in a previous article [13] on this subject which is similar to the present notation. However, it is planned to present a notation system elsewhere [19] to facilitate more comprehensive discussion of fundamental notions in complement-fixation reactions.

REFERENCES

1. Thompson, W. R. Use of moving averages and interpolation to estimate median-effective dose. 1. Fundamental formulas, estimation of error, and relation to other methods. *Bacteriological Reviews* 1947, *11*, 115–145.
2. Wadsworth, A. B. *Standard Methods of the Division of Laboratories and Research of the New York State Department of Health*; 3d ed. Baltimore, Williams and Wilkins, 1947, (a) p. 145–161.
3. Kärber, G. Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche. *Archive Experimentele Pathologie und Pharacologie*, 1931, *162*, 480–483.
4. Gaddum, J. H. Reports on biological standards. 3. Methods of biological assay depending on a quantal response. *Medical Research Council (Brit.), Special Report Series no. 183*, 1933, 46 p. H. M. Stationery Office, London.

5. Irwin, J. O. and Cheeseman, E. A. On an approximate method of determining the median effective dose and its error, in the case of a quantal response. *Journal of Hygiene*, 1939, *39*, 574–580.
6. Irwin, J. O. and Cheeseman, E. A. On the maximum-likelihood method of determining dosage-response curves and approximations to the median-effective dose, in cases of a quantal response. *Journal of the Royal Statistical Society* 1939 (Suppl.), *6*, 174–185.
7. Reed, L. J. and Muench, H. A simple method of estimating fifty per cent endpoints. *American Journal of Hygiene*, 1938, *27*, 493–497.
8. von Krogh, M. Colloidal chemistry and immunology. *Journal of Infectious Diseases*, 1916, *19*, 452–477.
9. Wadsworth, A., Maltaner, E., and Maltaner, F. The quantitative determination of the fixation of complement by immune serum and antigen. *Journal of Immunology*, 1931, *21*, 313–340.
10. Wadsworth, A., Maltaner, F., and Maltaner, E. Quantitative studies of the reaction of complement fixation with tuberculous immune serum and antigen. *Journal of Immunology*, 1938, *35*, 93–103.
11. Wadsworth, A. B. *Standard Methods of the Division of Laboratories and Research of the New York State Department of Health;* 2d ed. Baltimore, Williams and Wilkins, 1939, 681 p.
12. Maltaner, F. and Maltaner, E. The quantitative determination of the antigen-antibody reaction by complement fixation. *Third International Congress of Microbiology, New York. Report of Proceedings*, 1940, p. 781–782.
13. Thompson, W. R., and Maltaner, F. On the construction of graphs and tables for evaluation of the quantitative complement-fixation reactions and reaction ratios. *Journal of Immunology*, 1940, *38*, 147–157.
14. Rice, C. E. Studies of antipneumococcal serum. 2. Complement-fixing activity of antipneumococcal rabbit-serum with homologous type-specific carbohydrate. Technic of test. General quantitative relationships among reagents. *ibid.*, 1942, *43*, 129–148.
15. Mayer, M. M., Eaton, B. B., and Heidelberger, M. Spectrophotometric standardization of complement for fixation tests. *ibid.*, 1946, *53*, 31–35.
16. Kent, J. F., Bukantz, S. C., and Rein, C. R. Studies in complement fixation. 1. Spectrophotometric titration of complement; construction of graphs for direct determination of the 50 per cent hemolytic unit. *ibid.*, 1946, *53*, 37–50.
17. Rice, C. E. Studies of the complement-fixation reaction in virus systems. 1. Activities of vaccinia virus antigens and antisera. *ibid.*, 1946, *53*, 225–236; (a) p. 229.
18. Rice, C. E. Studies of the complement-fixation reaction in virus systems. 2. Activities of influenza virus antigens and antisera. *ibid.*, 1947, *56*, 343–356.
19. Thompson, W. R., Rice, C. E., Maltaner, E., and Maltaner, F. Some fundamental notions in estimation of complement fixation. 1. General relations and a proposed uniform notation. *Unpublished.*
20. Wilson, E. B. and Worcester, J. The determination of L.D. 50 and its sampling error in bio-assay. 1, 2, and 3. *Proceedings of the National Academy of Sciences, U. S.*, 1943, *29*, 79–85; 114–120; 257–262.
21. Worcester, J. and Wilson, E. B. A table determining L.D. 50 or the fifty per cent endpoint. *Proceedings of the National Academy of Sciences, U. S.*, 1943, *29*, 207–212.

22. Berkson, J. Application of the logistic function to bio-assay. *Journal of the American Statistical Association*, 1944, *39*, 357–365.
23. Berkson, J. Approximation of Chi-square by "probits" and by "logits." *Journal of the American Statistical Association*, 1946, *41*, 70–74.
24. Wilson, E. B. and Worcester, J. Bio-assay on a general curve. *Proceedings of the National Academy of Sciences, U. S.*, 1943, *29*, 150–154.
25. Winsor, C. P. A comparison of certain symmetrical growth curves. *Journal of the Washington Academy of Sciences*, 1932, *22*, 73–84.
26. Ipsen, J. *Contribution to the theory of biological standardization.* Nyt Nordisk Forlag, Arnold Busek, Copenhagen, 1941, 248 p.
27. Bliss, C. I. and Cattell, M. Biological Assay. *Annual Review of Physiology*, 1943, *5*, 479–539; (a) p. 493.
28. Northrop, J. H. The inactivation of trypsin. 3. Spontaneous inactivation. *Journal of General Physiology*, 1921–22, *4*, 261–274.
29. Hussey, R. G., and Northrop, J. H. A study of the equilibrium between the so called "antitrypsin" of the blood and trypsin. *Journal of General Physiology*, 1922–23, *5*, 335–351.
30. Granville, W. A., Smith, P. F., and Longley, W. R. *Elements of the differential and integral calculus.* New edition; New York, Ginn and Co., 1934, p. 380–381.
31. Cohen, Abraham. *An elementary treatise on differential equations.* New York, Heath and Co., 1906, 271 p.

QUERIES

QUERY: I am investigating the production of hard seeds in crim-
59 son clover. Since individual plants are extremely variable, I have devised a technique for splitting young plants—then growing the halves to maturity. The two half-plants vary much less than whole plants in their production of hard seeds.

The table shows the percent germination (that is, soft seeds) in half-plants grown in pots in the greenhouse. The arrangement of the pots was random. Each comparison consists of four pairs of half-plants. The third pair in Comparison II shows the typical variability of plants—in this respect, this experiment is the most uniform I have run.

PERCENT GERMINATION OF SEEDS FROM 16 PAIRS OF HALF PLANTS IN FOUR COMPARISONS

Comparison I		Comparison II		Comparison III		Comparison IV	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$c_1p_1k_1$	$c_2p_2k_2$	$c_1p_1k_2$	$c_2p_2k_1$	$c_1p_2k_1$	$c_2p_1k_2$	$c_1p_2k_2$	$c_2p_1k_1$
73	63	97	92	92	77	76	92
76	67	93	80	83	75	81	71
94	75	59	30	97	84	98	78
76	87	100	95	91	83	100	98
319	292	349	297	363	319	355	339

My main interest is in the main effects. If the high level of calcium leads to hard seed production, then all C_2 comparisons would be expected to have lower percent germinations. With the exception of two pairs, this was the case. The higher levels of phosphorous and potassium failed to produce such an effect.

I am not obviating the possibility of interactions. I will be very much interested in your method for obtaining these and in a test of significance of the effects.

ANSWER: These percentages are perhaps based on small and varying numbers of seeds. For exact work this should be investigated and suitable transformation made. Such refinements might affect the results given below but would not change the methods described.

Preliminary calculations indicate that the only effects that approach significance are the calcium main effect and the calcium-phosphorous interaction. The former is made up of the four differences:

$$\begin{aligned}
(8) - (1) &: c_2p_1k_1 - c_1p_1k_1 = 339 - 319 = 20\\
(6) - (3) &: c_2p_1k_2 - c_1p_1k_2 = 319 - 349 = -30\\
(4) - (5) &: c_2p_2k_1 - c_1p_2k_1 = 397 - 363 = -66\\
(2) - (7) &: c_2p_2k_2 - c_1p_2k_2 = 292 - 355 = \underline{-63}\\
&\qquad Total \qquad\qquad\qquad\qquad\quad -139
\end{aligned}$$

That is, calcium depressed the percent of germinating seeds, presumably stimulating the production of hard seeds. This stimulating effect of calcium was enhanced by the presence of phosphorous, the difference for computing the interaction being

$$(-66 - 63) - (-30 + 20) = -119$$

The two mean squares to be compared with the error variances are $(-139)^2/32 = 604$ and $(-119)^2/32 = 443$.

The two error variances involved, one between half-plants from the same plant and the other between the sums of these two half-plants, result in treatment comparisons at different levels of precision.

The error variances for the three main effects and for the second order interaction are the same, each being equal to the variance of half-plants from the same plant. To see this, observe that the variance for the calcium effect is, using the numerical designations of the treatments, the variance of

$$(8) + (6) + (4) + (2) - (1) - (3) - (5) - (7),$$

which may be written,

$$[(8) - (7)] + [(6) - (5)] + [(4) - (3)] + [(2) - (1)].$$

Each pair is made up of four differences between half-plants from the same plant. The variance of these differences is 133.24, double the mean square per half-plant shown in the analysis of variance. Since there are four differences per treatment pair and four such pairs, the variance of the total is 16(133.24) = 2,131.84; hence, the variance per half-plant is 2,131.84/32 = 66.6. There are three degrees of freedom for each of the four treatment pairs. For testing the significance of the calcium effect, we now have

$$F = 604/66.6 = 9.07, \qquad F_{.01} = 9.33.$$

The calcium-phosphorous interaction is given by

$$[(1) + (2)] + [(3) + (4)] - [(5) + (6)] - [(7) + (8)].$$

Here, as in the other two first order interactions, the treatment pairs are made up of four sums of half-plants, the variance of a sum being 782.6. Calculating as before, the variance per half-plant is now 391. F is $443/391 = 1.13$, obviously non-significant.

The analysis of variance of the entire experiment is shown in the table.

ANALYSIS OF VARIANCE OF PERCENT GERMINATION

Source of Variation	Degrees of Freedom	Mean Square
Treatment:		
C	1	603.78
P	1	11.28
K	1	0.28
CP	1	442.53
CK	1	69.03
PK	1	16.53
CPK	1	87.78
Differences between half-plants from the same plant	12	66.62
Sums of half-plants from the same plant	12	391.32

Had querist wished to evaluate all the effects with equal precision, he could have used balanced incomplete blocks with a pair of half-plants in each block. He has actually repeated the experiment with this design, and has been asked to present his results in "Queries".

This poses a nice problem for every experimenter: to foresee his needs with sufficient certainty to choose the exact design which will satisfy them.

George W. Snedecor

QUERY: I enclose data I have collected on chick growth and have
60 indicated what appears to be four possible ways of showing the significance of the differences between the means obtained. My problem is one of how best to present these data and their implications in print. Your comments and suggestions as to the correct usage of these four methods will be greatly appreciated.

In a feeding trial with chicks designed to study the comparative feeding values of six protein supplements, fourteen chicks were used per lot to begin with. The birds were all of the same sex and age and were from the same flock. During the course of the experiment, some chicks died due to unknown cause.

On the basis of chemical analyses, one half of the horsebean seed meal protein was replaced in the basal ration by the indicated protein supplements. All rations had the same level of crude protein.

The birds receiving each ration were housed together. Record was kept of the amounts of feed eaten by the lots only. The results are given in the table.

FINAL WEIGHTS OF CHICKS AT SIX WEEKS (GRAMS)

	Source of Protein Supplement					
	Horsebean	Linseed Oil Meal	Soybean Oil Meal	Sunflower Seed Oil Meal	Meat Meal	Casein
1	179	309	243	423	325	368
2	160	229	230	340	257	390
3	136	181	248	392	303	379
4	227	141	327	339	315	260
5	217	260	329	341	380	404
6	168	203	250	226	153	318
7	108	148	193	320	263	352
8	124	169	271	295	242	359
9	143	213	316	334	206	216
10	140	257	267	322	344	222
11		244	199	297	258	283
12		271	177	318		332
13			158			
14			248			
Mean	160.2	219.5	246.9	324.4	271.6	306.9

ANALYSIS OF VARIANCE

Source of Variation	Degrees of Freedom	Mean Square
Protein Supplements	5	43,627
Error	65	4,645

Four methods of comparison between lot means are proposed:

1. Fiducial limits for each lot mean.

2. Difference required for significance at 0.05 level for each individual pair of lots.

3. The largest value of the difference required for significance is taken as the common value for testing the significance for all pairs of lots.

4. For each individual pair of lots the value of t is used to test significance.

ANSWER: Your proposed method 1 is not a test of significance, but it is sometimes useful for assessing the results of an experiment like this if no set of independent comparisons is provided in the design.

Methods 2 and 4 produce results that are essentially the same. The 15 comparisons contemplated are not independent. The probabilities associated with them are not likely 0.05, and there is no way to determine what they are. For further comments on this method, see this Journal, Vol. 1, p. 26.

Method 3 is an inexact variant of method 2. The advantage is slight because all of the significant differences must be calculated in order to isolate the largest.

Subject to the severe limitations discussed below, I would suggest some such orthogonal set of comparisons as the following:

(a) Standard source (horsebean) vs. other sources.
(b) Vegetable sources vs. animal.
(c) Meat vs. casein.
(d) Two comparisons among vegetable sources.

Although your method of testing is inexact, this is a difficulty which I consider minor because decisions are usually based on many bits of evidence of which the test of significance is only one. Of far greater import are difficulties inherent in the design of the experiment. There is no suitable replication of the treatments. As a consequence, there are only ambiguous answers to questions supposed to have been cleared up.

Your estimate of error is based on the variation in weight of chicks housed together in the several lots. Two doubts assail such an estimate: (i) Are these weights independent or are there correlations due to environment? (ii) To what extent is the variation among these weights affected by differences in food intake? In the experimental design, no provision is made for answering either question.

In the treatment differences, possible effects of environment and food intake are confused with any real effects of the sources of protein. There is no way to untangle them. Some method of regression might be used to compensate for the differences in food consumed, but there is no comparable estimate of error.

You are to be commended for segregating the sexes for this experiment. I have encountered data in which confusion was worse confounded by unknown inequalities in the sex ratios.

It is clearly my opinion that questions about tests of significance are of little weight as compared to those about the design and conduct of the experiment. Unless unambiguous information is incorporated in the data statistical methods for extraction are futile. GEORGE W. SNEDECOR

THE BIOMETRIC SOCIETY

The Western North American Region of the Biometric Society was organized in Berkeley, California on June 24. Regional officers, nominated at the meeting and since elected by the Council, are

Vice-President, F. W. Weymouth, head of the department of physiology at Stanford University; Secretary-Treasurer, Mrs. Bernice Brown, statistical analyst at Project Rand of the Douglas Aircraft Co.; regional committee members, James E. Holloway, entomologist of the University of California, Jerome C. R. Li, assistant professor of mathematics at Oregon State College, G. A. Baker, assistant professor of mathematics and assistant statistician of the experiment station of the University of California, and Wade Rollins, research assistant at the University of California, College of Agriculture at Davis.

The Australian Region is well along in formation with E. A. Cornish and Helen N. Turner carrying the ball. A Regional meeting is planned at Hobart in January, in conjunction with the meeting of the Australian and New Zealand Association for the Advancement of Science. Permanent officers will be named at that time, and decisions made as to future policy and regional meetings.

Several New Zealanders, interested in the Society, have proposed organizing a New Zealand Region because of their distance from Australia. Dr. J. T. Campbell, senior lecturer in mathematics at Victoria University in Wellington, has arranged a gathering of New Zealand statistical workers at the University in late August and plans to bring up the question at that time.

The Eastern North American Region is holding a joint session with the Vital Statistics Section of the American Public Health Association on Wednesday, November 10 during the annual meeting of the Association in Boston.

A round table discussion of Morbidity Surveys is scheduled for 9:30 A.M. in the Paul Revere Banquet Hall, Mechanics Building. General discussion will follow the panel participants.

Moderator: Mr. George St. J. Perrott, Chief, Division of Public Health Methods, U. S. Public Health Service

Public Health Statistician: Mr. Theodore D. Woolsey, Division of Public Health Methods, U. S. Public Health Service

Sampling Statistician: Mr. Nathan Keyfitz, Mathematical Advisor, Dominion Bureau of Statistics, Ottawa, Canada

Opinion Poll Statistician: Dr. George Gallup, Director, American Institute of Public Opinion, Princeton, New Jersey

Epidemiologist: Dr. Alexander D. Langmuir, Associate Professor, Department of Epidemiology, School of Hygiene and Public Health, Johns Hopkins University

Health Officer: Dr. Huntington Williams, Commissioner of Health, City of Baltimore

BIOMETRY

R. A. Fisher*

The rise of biometry in this 20th century, like that of geometry in the 3rd century before Christ, seems to mark out one of the great ages or critical periods in the advance of the human understanding. From its humble beginnings in meeting the needs and satisfying the practical requirements of the work of land measurement and architecture, *geometry* found its way, by the enchanting clarity of its concepts and processes, into the heart of what the Greek world meant by a liberal education; an education that is fit for *free* men who would think for themselves, and not fit only for slaves and officials whose aims and concepts were dictated from above. It was the liberation of the spirit experienced by the Greek students of geometry which gave the subject to them the exalted status it undoubtedly held, and won the veneration of the entire period. We can, I think, partly understand their feeling, when we realise that here for the first time the human spirit came to handle abstractions, of their nature necessarily timeless and perfect, and to handle them with confidence, because they were well defined. The well defined abstraction seems, in fact, to be the invention of the Greek geometers, and an invention of lasting significance to human thought.

But it was not merely its conceptual clarity which gave to geometry its fascination. With well defined concepts the intellect found itself capable of acting with unprecedented efficiency. Men learnt to reason, deductively, from well defined abstract concepts, to cogent and irrefrageable conclusions. And with its use, with its exercise, in the field of geometry, the *principles* of deductive reasoning came to be understood, or at least, to be codified, so as to give rise to the subject known as *Logic*. It has been a fashion among some modern mathematicians to speak of Mathematics itself, or themselves, as but a branch of logic. This, of course, is but a formalisation, appropriate to a purely deductive habit of mind. The historical fact unquestionably shows logic as a later growth, a formulation of the thought processes, in which the practice of geometry had already made man sufficiently adept, to ensure agreement as to general principles. And the conclusions of geometers themselves, apart from the artistically unified presentation which Euclid gave them, embody the horse-sense of ages of predecessors trying to measure accurately, and using increasingly subtle and indirect means of measurement and of accurate construction. Even in Euclid's treatise it is discernible that the grand aim towards which the whole edifice of theorems and problems is directed is the practical and exact construction of the five perfect or absolutely regular figures which are possible in three dimensions.

---

*R. A. Fisher, president of the Society, addressed the Inaugural meeting of the British Region in London on April 29. His address was such an adroit summary of the reasons for launching a new scientific Society that we are publishing it in this issue of BIOMETRICS.

Now, I suppose circumstances might have conspired to give to surveying, or to astronomy, or to any other subject sufficiently rich in observational detail, the honour of compassing the second great stage of intellectual liberation, by making known the principles of that second and scarcely explored mode of logic, which we know as induction; of clarifying the principles of reasoning from the particular to the general, from the observations to the hypotheses, in ways necessarily inaccessible to purely deductive logic, or to any mathematics which can properly be regarded as derivable wholly from deductive logic, of making men free to recognize with certainty the consequence not of axioms or dogmas, but of carefully ascertained facts. But, as it has happened, it has been reserved for Biometry, the active pursuit of biological knowledge by quantitative methods, to take this great step; and the man who in the nineteenth century did more than any other to prepare the way was, I think, undoubtedly Francis Galton.

The peculiarity of Galton's temperamental make-up which led him to play this part was, in my opinion, the insistent need that he felt to think constructively about variable phenomena. Unquestionably he was led to concentrate his attention upon variation, through the central place which variation held in the theory of evolution, which his half-cousin Charles Darwin had put forward, and which influenced Galton profoundly, as appears clearly in his book *Hereditary Genius*, published after the *Origin* by only ten years. To Galton, however, variation of all kinds had an appeal, or a fascination, as much in meteorology for example as in heredity, and this appeal we can appreciate if we consider what an obstacle to coherent thought mere quantitative variation had formerly been. Even now, common phrases and modes of thought express this impotence from which Galton's generation was just emerging, thanks largely to Galton's own efforts. If one were to say: "Nothing definite can be asserted about the political opinions of entomologists, for their opinions vary", even an audience of biometers might admit the statement as rational, although they all know perfectly well, from their own constant experience, that a great variety of definite statements could be made about every variable phenomenon that had been studied. Without this experience, however,—and the bulk of mankind are without this experience,—without the modern concept of frequency distributions, and the habit of thinking coherently in terms of frequency distributions, thought comes to a full stop. The urge, apparent in elementary text books of statistics, to find "measures of central tendency" as conceptually constant substitutes for the really variable values, embalms one of the earliest efforts to evade the intellectual difficulty. With better apparatus of thought at our disposal, we can now reinterpret the measures as statistical estimates. As such they make sense; but we should remember that they are still introduced and taught at a stage when there is as yet no thought of estimation. Then again, we still have the administrative compromise, such as "A fair wage is one sufficient to maintain in decency a wife and three children", and it is with pained surprise, and with great reluctance, that the administrator learns to admit that such a decision will leave half the real children of the country, belonging to families of four or more, insufficiently provided for, and that at the same time it saddles the wage fund, and therefore the purchaser of goods, with providing for about twenty million non-existent children. In this aspect family allowances constitute an elementary recognition of the unwelcome fact of biological variability.

The primitive function of the biometric movement, characteristics of the present century, is therefore to conserve by constant use, and incidentally to improve and refine, the thought forms, which make possible an understanding of variable phenomena. These phenomena come to our knowledge by observation of the real world, and it is no small part of our task to understand, design and execute the forms of observation, surveys or experiments, which shall be competent to supply the knowledge needed. The observational material requires interpretation and analysis, and no progress is to be expected without constant experience in analysing and interpreting observational data of the most diverse types. Only so, as I have suggested, can a genuine and comprehensible formulation of the processes of inductive reasoning come into existence. As we bear these objects in mind, as we allow ourselves to appreciate their immense practical importance, as we yield to their intellectual fascination, so, it is common experience, we come to think of ourselves less in terms of the special scientific disciplines, less as chemists or entomologists, geneticists or mathematicians, and more in terms of the community of our interests with those doing similar work in other departments. It is to promote interchange of ideas, personal contacts, and mutual appreciation of our diverse problems and methods, that we have felt the need of a new scientific organisation, in which our work may be viewed in a new perspective, not as something extraneous and eccentric, a funny sort of botany, for example, or of palaeontology, or of medicine, but as a tidal movement of our time, which has already begun to refresh and reinforce the means of research in all the biological sciences.

NEWS AND NOTES

CANADA—**James W. Fisher**, Virologist, Laboratory of Hygiene, Department of National Health and Welfare, Ottawa, states that they are happy in their new quarters. In response to an inquiry regarding the use of statistical tools at the Laboratory, he replies: "The Laboratory of Hygiene, being the control laboratory for certain biologicals designed for human use in Canada, finds statistical tools to be of inestimable value especially for assay work. The graded method has found wide application in the Antibiotic Section; while the quantal has justified its worth in assaying tetanus and diphtheria toxoids. The aid of the statistician has been welcomed in the Bacteriology Section where problems are encountered that deal with the quantitative estimation of the number of microbes in foods, or surviving in the presence of various disinfectants. In estimating certain properties of viruses by their lethal or other effects produced in animals or eggs, under various experimental conditions, statistical tools have been found to be potent weapons—the 'methodology' of the analysis of variance in particular. Those in the Virus Section are becoming familiar with probits, standard errors, tests of significance, etc., when referring to either *in vitro* or *in vivo* laboratory procedures. We are using and attempting to develop better methods so that our experimental procedures will be more efficient and our results mathematically sound." . . . **Allan Paull** completed his Ph.D. in statistics and has returned to the Grain Research Laboratory, Board of Grain Commissioners, Winnipeg, Manitoba. His dissertation was "On A Preliminary Test for Pooling Mean Squares in the Analysis of Variance," under the supervision of **W. G. Cochran.**

CHINA—**C. M. Wang** has, at our request, sent us information regarding The Biometry Laboratory in the College of Agriculture, Taiwan National University. The Laboratory was established in the fall of 1946 "(1) for teaching college students and training junior staffs of other laboratories and research stations in Taiwan (Formosa) in order to rationalise the experiments, to save time and expense in the interval from test to practice and thus to prompt and to facilitate the agricultural extension, (2) to publish the results of experiments and researches." They have paid attention especially to the statistical methods of small samples and the consistency of practice with the underlying theory. The staff consists of Director Wang, two lecturers, two assistant teachers and seven other assistants. Their curriculum includes courses in biometry, mathematics of statistics, field plot techniques and special lectures on problems of biometrical technique. "Study on correlations among various characteristics of rice plants since the spring of 1947 and field technique research in rice experiments since the spring of current year are being conducted. Experiments with rice, jute and sugar cane are in progress cooperating with the Taiwan Province College of Agriculture, the Taiwan Agricultural Research Station and the Sugar Cane Research Institute of Taiwan Sugar Corporation."

ENGLAND—**John Wishart** is to be a Visiting Professor at the University of North Carolina, Institute of Statistics, Raleigh, during the spring quarter of 1949. He will give a course in Experimental Designs. Dr. Wishart has been interested in biometric statistics for many years, having assisted Karl Pearson and R. A. Fisher in their laboratories before going to the School of Agriculture, Cambridge.

INDIA—The Statistics Section of the Indian Science Congress met this year at Patna. A symposium on "crop-forecasts from a study of weather conditions" was held. **L. A. Ramdas,** Agricultural Meteorologist of Poona Meteorological Observatory, spoke of his experiences about the extent of variation in crop-yields due to weather. The other speakers discussed the effects of specific meteorological factors on various crops . . . **U. S. Nair,** Professor of Statistics, Travancore University, was elected the President; **Sadasiv Sengupta,** Statistical Officer of the East India Railway, Recorder of this Statistics Section . . . In the March 1948 bulletin of the Calcutta Statistical Association, **K. K. Mathen,** All-India Institute of Hygiene and Public Health, Calcutta has an expository article on "Studies on the sampling procedure for a general health survey" . . . A statistical conference was held January 27 to 31, 1948, at Singapore. It was presided over by **W. Clyde.** The countries participating were: Australia, Burma, French Indo-China, Malaya and British North Borneo, Hongkong, India, The Malaya Union, The Netherlands East Indies, Siam and Singapore. India was represented by **P. V. Sukhatme** and **G. M. Sankpal.** The task before the conference was essentially of an exploratory character aimed at standardizing rice crop statistics and improving the methods of collecting them.

ITALY—**F. Brambilla**, of the University "Bocconi" in Milan, held last winter a course on "Statistics for Geneticists" at the Zoological Station in Naples. The course was attended by members of the staff of the Zoological Station, guests and students of the University of Naples . . . **L. L. Cavalli**, of Milan, recipient of a fellowship from the Italian National Research Council, is working at the John Innes Horticultural Institution, London, with **K. Mather** on problems of biometry and statistical genetics.

PALESTINE—**Jos. Carmin,** Independent Biological Laboratories, Kefar-Malal, P. O. Ramatayim writes: "As to statistical aspects of our research, I am trying my best to have it up to date. In planning any research work the statistical aspect is always taken into consideration as well as in working out the data."

UNITED STATES—The Statistical Section of the Tennessee Public Health Association met May 4, at Nashville, Tennessee. **Ann Dillon,** chairman, presided. At the morning session papers were given by **Eleanor Gorham,** Nashville Council of Community Agencies; **Parker Mauldin,** Division of Medical Research Statistics, Veterans Administration; **Jack Moshman,** Office of Medical Advisor, Atomic Energy Commission, Oak Ridge; and **Helen C. Maher,** Tennessee Valley Authority. After a luncheon meeting, the speakers at the afternoon session were **Carolina Randolph,** The Commonwealth Fund; **William F. Elkin,** Oak Ridge Health Department; **Sara Lou Hatcher,** Tennessee Department of Public Health; and **Mildred Patterson,** Sumner County Health Department. At a business meeting, the following officers were elected for 1948-1949: Chairman, **Margaret Martin,** Vanderbilt University, Department of Preventive Medicine and Public Health; Vice-Chairman, William F. Elkin; and Secretary, Sara Lou Hatcher . . . **Alexander M. Mood** and **Arthur J. Brown,** members of the staff of the Statistical Laboratory, Iowa State College, are on leave with Project Rand, Santa Monica, California . . . The Virginia Academy of Science meetings were held May 6 to 8, at Hotel Roanoke. At the sessions of the Section on Statistics, **A. E. Brandt** gave talks on "Some statistical aspects of experimentation" and "On the correlation of a part with the whole." The officers for this section are **Boyd Harshbarger,** Chairman; **W. H. White,** Vice-Chairman; and **Walter Hendricks,** Secretary. Mr. Harshbarger was also selected as President-Elect for the Virginia Academy of Science . . . **J. Neyman** reports the following appointments and promotions in the Statistical Laboratory, University of California: **Michel Loeve,** formerly Reader in Mathematics, Kirkwood College, University of London, appointed Professor of Mathematics and Research Associate in the Statistical Laboratory; **Charles M. Stein** promoted to Assistant Professor and Research Associate; **Elizabeth L. Scott** promoted to Lecturer and Research Assistant; **Edith Mourier,** formerly Research Assistant at the Institute Henri Poincaré, Paris, appointed Research Assistant and Teaching Assistant at the Statistical Laboratory; and **David Rubinstein** promoted to Teaching Assistant.